



DÉPARTEMENT DES SCIENCES BIOLOGIQUES

THÈSE DE DOCTORAT EN BIOLOGIE

Spécialité : Génétique & Biologie Moléculaire

Présentée par :

Mme. Amel OUESLATI BAHRI

Intitulée :

Diversité et origine génétique des espèces de la sous famille des Aurantioideae et décryptage des structures interspécifiques des génomes des agrumes modernes

Membres du Jury :

Président	Mr Aly RAÏES,	Professeur, FST
Directeur de Thèse	Mme Amel HANNACHI SALHI,	Professeur, FST
Rapporteur	Mr Mohamed EL GAZZAH,	Professeur, FST
Rapporteur	Mme Nejia ZOGHLAMI,	Maître de conférences, CBBC
Examineur	Mme Faten GORSAN FAKHFAKH,	Maître de conférences, FSB
Invité	Mr Patrick OLLITRAULT,	Directeur de recherche, CIRAD-UMR-AGAP

Année Universitaire 2016 / 2017

Dédicaces

Faire de la recherche scientifique est un projet auquel j'ai aspiré encore très jeune.

Ce rêve de jeune fille a été approfondi jour après jour durant mon cursus académique. J'ai eu la chance d'avoir une éducation scientifique et littéraire.

Nourrie et influencée par la littérature et par la philosophie scientifique, des questions et des réflexions ont commencé à s'installer en moi. Elles ont contribué à ma formation personnelle.

La biologie, quant à elle, était ma passion, par la richesse de la nature et de sa diversité, elle a ainsi suscité ma curiosité. L'immense richesse de cette nature par sa faune et sa flore, de cette planète, de cet univers et de son évolution au cours du temps m'ont vraiment toujours impressionné et m'ont confirmé que nous n'en sachons qu'une chose très minime. La nature agit, s'adapte, sélectionne et évolue.

La vérité hypothétique des sciences expérimentales n'a cessé de montrer la puissance de sa démarche quant à la production de connaissances valides conformes aux faits et à leurs relations testables de cause à effet. Les sciences depuis l'antiquité n'ont pas cessé d'évoluer et de répondre aux questions relatives à ce monde au comment et au pourquoi. Une vérité que je trouve absolue, c'est que plus on réunit de connaissances, plus on est capable de juger qu'on en possèdera qu'imparfaitement.

Le choix de la génétique est venu après. Il correspondait à la cohérence de la logique des maths et de la richesse de la biologie, un mariage de la rigueur du raisonnement logique et de la curiosité suscitée par la diversité de la nature.

Voir un rêve se réaliser est un pur bonheur, c'est avec ces mots enthousiastes que je décris mon immense honneur et gratitude. Durant ces cinq ans de thèse, je me suis menée en bateau à une aventure où j'ai rencontré des hauts et des bas. Cette expérience est plus qu'un travail de recherche, elle m'a marqué sur tous les niveaux certainement professionnel et aussi relationnel et humain où le développement personnel était mon plus grand gain. J'ai évolué judicieusement à ses obstacles grâce au soutien de ma famille qui a cru en moi et m'a poussé à positiver sans être crédule. Mes parents n'ont pas cessé de m'encourager, me

soutenir, me réconforter et surtout m'aimer durant mon parcours d'étude de l'enseignement primaire à l'universitaire. Cette thèse ; l'œuvre de mes parents, je la leur dédie avec un pur remerciement et reconnaissance pleins d'amour.

Je la dédie à mon frère Nidhal pour tout instant de joie, je lui souhaite le bonheur et la réussite.

Une dédicace spéciale à mon très cher mari Karim pour son soutien, son encouragement et sa présence tout au long de ces dernières années de finalisation de la thèse. Que ce travail témoigne de ma reconnaissance et de mon amour sincère et fidèle.

Remerciement

*Le travail présenté dans cette thèse a été effectué au sein du laboratoire de Génétique Moléculaire, Immunologie et Biotechnologie du Département des Sciences Biologiques à la Faculté des Sciences de Tunis, Université de Tunis El Manar, sous la direction bienveillante de Madame **Amel HANNACHI SALHI**.*

*Des stages ont été effectués durant cette thèse au sein du laboratoire de Protection Végétale et Biotechnologie du centre de recherche Insituto Valenciano de Investigaciones Agrarias sous la direction de Monsieur **Patrick OLLITRAULT**.*

Ce travail n'aurait pas pu voir le jour sans la collaboration sérieuse et l'aide précieuse de plusieurs personnes. J'espère trouver les mots justes pour leur témoigner toute ma reconnaissance et ma gratitude.

*Je voudrais adresser mes sincères remerciements à Madame **Amel HANNACHI SALHI**, Professeur à la Faculté des Sciences de Tunis et directrice de ma thèse de doctorat. Vos conseils et vos précieuses directives, toujours donnés avec une extrême amabilité, m'ont permis d'élaborer cette thèse. Vos compétences et votre bienveillance méritent toute admiration. En considération de votre savoir et de vos qualités humaines, veuillez trouver dans ce travail l'expression de ma profonde gratitude et mon grand respect.*

*Je tiens à présenter mes gratitudes les plus sincères à Monsieur **Patrick OLLITRAULT**, Directeur de recherche au Centre de Coopération Internationale en Recherche Agronomique pour le Développement et responsable de mes stages. Je vous remercie pour la confiance que vous m'avez témoignée en m'acceptant au sein de votre équipe. Merci pour votre accueil chaleureux et vos précieux conseils. Merci d'avoir réuni les conditions nécessaires au bon déroulement de mes travaux. A l'égard de votre dévouement, je vous exprime mon profond respect et mon immense reconnaissance.*

*Je tiens à remercier vivement Madame **Neila FARAH TRIFI**, Professeur et Directrice du Laboratoire de Génétique moléculaire, Immunologie et*

Biotechnologie à la Faculté des Sciences de Tunis, pour m'avoir accueilli au sein du Laboratoire de Génétique moléculaire, Immunologie et Biotechnologie et pour la confiance. Je vous prie de croire en ma profonde estime et à ma haute considération.

*Que Monsieur **Aly RAÏES**, Professeur à la Faculté des Sciences de Tunis, trouve ici l'expression de ma profonde gratitude pour la disponibilité dont il a fait preuve et pour m'avoir fait l'honneur d'accepter de présider le jury de ma thèse.*

*Je remercie vivement Monsieur **Mohamed El GAZZAH**, Professeur à la Faculté des Sciences de Tunis, pour le temps qu'il a accordé en prenant en charge d'accepter d'être rapporteur de cette thèse.*

*J'adresse mes plus vifs remerciements à Madame **Nejia ZOUGHLAMI**, Maître de conférences au Centre de Biotechnologie de Borj Cedria, de l'intérêt qu'elle a manifesté à l'égard de cette recherche en s'engageant à être rapporteur.*

*Je remercie vivement Madame **Faten GORSANE FAKHFAKH**, Maître de conférences à la Faculté des Sciences de Bizerte, pour ses conseils et son soutien continu et pour avoir accepté d'examiner ce travail.*

*Je remercie vivement Monsieur **Hatem FAKHFAKH**, professeur à la Faculté des Sciences de Bizerte, pour son soutien inconditionnel et sa disponibilité.*

*Je tiens à remercier vivement Monsieur **Luis NAVARRO**, Professeur et Directeur de Recherche à Instituto Valenciano de Investigaciones Agrarias, qui a eu la gentillesse de m'accepter au sein de son honorable institution.*

*Je tiens à remercier Madame **Ghada BARAKET**, Maître assistante à l'Institut Préparatoire aux Etudes d'ingénieurs de Bizerte, pour l'aide et l'attention qu'elle a consacré à mon égard.*

*Je remercie de même Madame **Salwa ZEHDI**, Maître assistante à la Faculté des Sciences de Tunis, pour son encouragement continu.*

*Ces remerciements seraient incomplets si je n'en adressais pas à l'ensemble des **MEMBRES** du Laboratoire de Génétique moléculaire, Immunologie et Biotechnologie et surtout à l'équipe 'Arbres fruitiers' pour leur soutien et pour la très bonne ambiance que j'ai toujours trouvée. Je remercie plus particulièrement*

Donia pour nos nombreuses discussions et nos moments spéciaux partagés. Je voudrais aussi remercier Aymen.

J'adresse mes plus vifs remerciements à tous les membres du laboratoire CIRAD-IVIA, pour leur bonne humeur. Je remercie plus particulièrement Madame Frédérique OLLITRAULT et Monsieur Pierre Mournet pour leur aide et leur disponibilité.

Sommaire

Introduction Générale	1
Etude Bibliographique	5
1. Origine géographique et histoire des Agrumes	5
2. Etude Botanique et description morphologique	7
2.1 La famille des Aurantioideae.....	7
2.2 Le groupe des Agrumes vrais	9
2.3 Le genre <i>Citrus</i> L.....	11
2.3.1 Origine génétique des taxons de base.....	12
2.3.2 Origine génétique des variétés secondaires	14
3. L'information génétique chez les agrumes : Structure, Ploïdie et Mode de transmission	21
4. Importance économique et distribution géographique	22
4.1 A l'échelle mondiale	22
4.2 A l'échelle nationale	23
5. Les principales contraintes de l'agrumiculture	26
6. L'amélioration des agrumes.....	27
6.1 Principaux objectifs	27
6.2 Atouts et facteurs limitants	28
7. Application des marqueurs génétiques chez les Agrumes	29
7.1 Evolutions des techniques moléculaires et étude de la diversité génétique.	30
7.2 Applications à l'étude des taux d'apomixie.....	32
7.3 Application aux études phylogénétiques	33
7.3.1 Phylogénie chloroplastique	34
7.3.2 Phylogénie nucléaire.....	35
7.4 Application à la cartographie génétique	36
7.5 Application à l'étude du déterminisme génétique des caractères (QTL)	37
8. Progrès récents dans le domaine de la génomique chez les Agrumes.....	37
8.1 Les séquences de référence complète du génome.....	37
8.2 Re-séquençage et décryptage des structures génomiques interspécifiques	40
8.3 Puce à ADNs et étude de l'expression du génome	41
8.4 Les promesses de la GBS pour un décryptage des structures génomiques étendues à l'ensemble des collections et des populations recombinantes	41

Matériel & Méthodes.....	44
1. Matériel végétal et séquences utilisées	45
1.1 Analyse de l'espaceur intergénique <i>trnL</i> _(UAA) - <i>trnF</i> _(GAA)	45
1.2 Analyse basée sur les marqueurs moléculaires SNPs associée au génotypage par la PCR compétitive allèle-spécifique (KASPar)	46
1.2.1 Sélection des séquences chloroplastiques	46
1.2.2 Matériel végétal utilisé pour l'analyse des données KASPar	46
1.3 Analyse basé sur la Technique de Génotypage par Séquençage (GBS)...	49
2. Etude du polymorphisme moléculaire et méthodes utilisées	51
2.1 Amplification de l'espaceur intergénique <i>trnL</i> _(UAA) - <i>trnF</i> _(GAA)	51
2.2 Le séquençage selon la technique de Sanger	52
2.3 Génotypage par la méthode KASPar™	52
2.4 La technique de génotypage par séquençage (GBS : Genotyping By Sequencing)	53
2.4.1 Préparation des librairies et séquençage	53
2.4.2 L'appel des génotypes SNPs et InDels	54
3. Analyse <i>IN SILICO</i> des séquences de l'espaceur intergénique <i>trnL</i>_(UAA)- <i>trnF</i>_(GAA)	54
3.1 Vérification de l'identité des séquences	54
3.2 Visualisation, correction et alignement des séquences	55
3.3 Identification et annotation des séquences du pseudogène <i>trnF</i> _(GAA)	55
3.4 Taille, composition nucléotidique et identification des différents types de mutations au niveau des séquences étudiées	56
4. Identification et sélection des SNPs	56
5. Paramètres génétiques	57
5.1 Diversité haplotypique	57
5.2 Diversité nucléotidique	57
5.3 Le contenu informatif du polymorphisme (PIC)	57
5.4 La fréquence allélique F	58
5.5 Le taux d'hétérozygotie observé et attendue	58
5.6 Indice de fixation F_{IS} de Wright	58
5.7 Analyse de la différenciation génétique	59
5.7.1 Indice de différenciation génétique F_{ST}	59
5.7.2 Indice de différenciation génétique G_{ST}	59
5.7.2.1 Méthode utilisée pour l'analyse des données KASPar	60
5.7.2.2 Méthode utilisée pour l'analyse des données GBS	60
5.7.3 Le coefficient de différenciation génétique NST	61
5.7.4 Flux de gènes et nombre de migrants efficaces N_m	61
5.8 Analyse cladistique	61
5.8.1 La méthode du plus proche voisin (NJ : Neighbor-joining)	61
5.8.2 La méthode de regroupement non pondéré par paire avec moyenne arithmétique : UPGMA	62

5.9 Analyse factorielle sur tableau de distances	63
6. Etude de l'évolution moléculaire	63
6.1 Test de Tajima	63
6.2 Test de Fu & Li.....	64
6.3 Test de neutralité de Fu	64
Résultats & Discussions.....	65
Chapitre I : Etude du polymorphisme génétique, de la phylogénie et de l'évolution moléculaire de l'espaceur intergénique <i>trnL</i>_(UAA)-<i>trnF</i>_(GAA) de l'ADN chloroplastique chez les espèces tunisiennes du genre <i>Citrus</i>	66
1. Amplification de l'espaceur intergénique <i>trnL</i>_(UAA)-<i>trnF</i>_(GAA).....	67
2. Séquençage de l'espaceur intergénique <i>trnL</i>_(UAA)-<i>trnF</i>_(GAA) de la région non codante chloroplastique	67
2.1 Vérification de l'identité des séquences par BLAST	67
2.2 Analyse <i>in silico</i> des séquences de l'espaceur intergénique <i>trnL</i> _(UAA) - <i>trnF</i> _(GAA)	69
2.2.1 Variation de la longueur et de la composition nucléotidique des séquences <i>trnL</i> _(UAA) - <i>trnF</i> _(GAA)	69
2.2.2 Diversité génétique	70
2.2.3 Analyse phylogénétique de l'espaceur <i>trnL</i> _(UAA) - <i>trnF</i> _(GAA)	71
2.2.4 Analyse du pseudogène <i>trnF</i> _(GAA)	75
2.3 Evolution moléculaire	78
2.3.1 Les tests de Tajima's et de Fu & Li.....	78
2.3.2 La statistique de Fu.....	78
2.3.3 Analyse de la différenciation génétique	79
3. Conclusion.....	80
Chapitre II : Elaboration d'une clé moléculaire taxonomique de la sous famille des Aurantioideae en utilisant les marqueurs moléculaires SNPs diagnostiques des principaux clades génotypés par la PCR compétitive allèle-spécifique	82
1. Identification des SNPs des clades spécifiques.....	83
2. Analyse KASPar	89
3. Utilité et complémentarité de l'analyse KASPar/SNPs diagnostiques par rapport à d'autres marqueurs moléculaires	94
4. Conclusion.....	96

Chapitre III : Révélation des structures interspécifiques <i>C. maxima</i>/<i>C. reticulata</i> au niveau des génomes des variétés de mandariniers modernes et des hybrides de mandariniers par la technique du Génomage par Séquençage (GBS).....	98
1. Recherche des marqueurs diagnostiques de la différenciation de <i>C. maxima</i> / <i>C. reticulata</i>	99
2. Appel des génotypes et diversité variétale	100
3. Diversité entre les mandarines et les pamplemousses et recherche de polymorphismes diagnostiques de la différenciation de <i>C. reticulata</i> / <i>C. maxima</i>	101
4. Introgression interspécifique chez les variétés représentatives des mandarines et des pamplemousses	103
5. Structure phylogénomique des variétés modernes	107
5.1 Validation de l'approche GBS pour décrypter les structures phylogénomiques par l'analyse de quatre variétés précédemment analysées par WGS (Wu <i>et al.</i> , 2014).....	109
5.2 Les structures phylogénomiques des variétés représentatives des taxons de base <i>C. reticulata</i> et <i>C. maxima</i>	111
5.3 Les structures phylogénomiques des mandarines hybrides, tangors, tangelos, pomelos et orangelos	114
6. Conclusion.....	121
Discussion et Conclusions générales & Perspectives.....	124
Références Bibliographiques	131
Annexes.....	156

Liste des Abréviations

ADN :	Acide DésoxyriboNucléique
ADNcp :	ADN ChloroPlastique
ADNr :	ADN Ribosomal
AFLP :	Amplification Fragment Length Polymorphism
ARS :	Agricultural Research Service
BAC :	Bacterial Artificial Chromosome
BLAST :	Basic Local Alignment Search Tool
CAPS :	Cleaved Amplified Polymorphic Sequence
CIRAD :	Centre de Coopération Internationale en Recherche Agronomique pour le Développement
CLAM :	Comité de Liaison des Agrumes Méditerranéenne
CRB :	Center of Biological Ressources
CTA :	Centre Technique des Agrumes
CTV :	Citrus Tristeza Virus
DAMD :	Directed Amplification of Minisatellite-region DNA
ddNTP :	didésoxyNucléotides TriphosPhate
dNTP :	désoxyNucléotides TriphosPhate
EDTA :	EthyleneDinitro Tetraacetic Acid
EST :	Expressed Sequence Tags
FAO :	Food and Agriculture Organisation
FRET :	Fluorescent Resonance Energy Transfert
g :	Gramme
GBS :	Genotyping By Sequencing
GIF :	Groupement Interprofessionnel des Fruits
HSVd :	Hop stunt viroid
ICGC :	International Citrus Genome Consortium
Indel :	Insertion/ Délétion
INRA :	Institut National de la Recherche Agronomique
IR :	Inversée Répétée
IRAP :	Inter-Retrotransposon Amplified Polymorphism
ISSR :	Inter Simple Sequence Repeats
ITS :	Internal Transcribed Spacer

IVIA : Instituto Valenciano de Investigaciones Agrarias
JC : Jésus-Christ
Kb : Kilobase
LSC : Large Simple Copie
MAS : Marker Assisted Selection
Mb : Mégabase
mM : Millimolaire
NCBI : National Center of Biotechnology Information
ng : Nanogramme
NGS : Next Generation Sequencing
NJ : Neighbor-Joining
nm : Nanomètre
Pb : Paire de bases
PCR : Polymerase Chain Reaction
QTL : Quantitative Trait Loci
RAPD : Random Amplification of Polymorphic DNA
RFLP : Restriction Fragment Length Polymorphism
SCAR : Sequence-Characterized Amplified Region
SCoT : Start Codon Targetted
SNP : Single Nucleotide Polymorphism
SRA : Station de Recherches Agronomiques
SRAP : Sequence Related Amplified Polymorphism
SSC : Simple Seule Copie
SSCP : Single Strand Conformation Polymorphism
SSR : Simple Sequence Repeats
Tm : Température d'hybridation
UPGMA : Unweighted Pair Group Method with the Arithmetic Average
USDA : United States Department of Agriculture
µg : Microgramme
µL : Microlitre

Liste des Figures

Figure 1 : Carte représentative de l'origine et de la diffusion des agrumes dans le monde (Praloran, 1971)	6
Figure 2 : Taxonomie de la sous famille des Aurantoideae (Swingle & Reece, 1967).....	8
Figure 3 : Fruit de l'espèce <i>Fortunella margarita</i> 'Nagami kumquat' (Aujardin, 2016)	9
Figure 4 : Fruit de l'espèce <i>Poncirus trifoliata</i> (Thibault, 2016)	9
Figure 5 : Fruit de l'espèce <i>Eremocitrus glauca</i> (Lucas, 2015)	10
Figure 6 : Fruit de l'espèce <i>Microcitrus australisica</i> 'Citron caviar' (a), Vésicules d'un fruit 'Citron caviar' coupé en deux (b) (Clémentine, 2012)	10
Figure 7 : Fruit de l'espèce <i>Clymenia polyandra</i> (Wakonai, 2016)	11
Figure 8 : Fruit de l'espèce <i>Citrus maxima</i> (Berger, 2003)	13
Figure 9 : Fruit de l'espèce <i>Citrus medica</i> 'Main de Bouddha' (Ly, 2014).....	13
Figure 10 : Fruit de l'espèce <i>Citrus reticulata</i> (Amel Oueslati, 2010).....	14
Figure 11 : Fruit de l'espèce <i>Citrus micrantha</i> (Robyn Jay, 2016).....	14
Figure 12 : Représentation schématique de l'origine de quelques espèces secondaires du genre <i>Citrus</i> (Ollitrault <i>et al.</i> , 2012 avec modification par Amel Oueslati).....	15
Figure 13 : Fruit de l'espèce <i>Citrus aurantium</i> (GIF, 2010)	16
Figure 14 : Fruit de l'espèce <i>Citrus sinensis</i> 'Maltaise' (GIF, 2010).....	17
Figure 15 : Fruit de l'espèce <i>Citrus limon</i> (GIF, 2010).....	17
Figure 16 : Fruit de l'espèce <i>Citrus paradisi</i> (GIF, 2010).....	18
Figure 17 : Fruit de l'espèce <i>Citrus aurantifolia</i> (Aujardin, 2016).....	19
Figure 18 : Fruit de l'espèce <i>Citrus clementina</i> 'Marisol' (GIF, 2008)	19
Figure 19 : Fruit du groupe horticultural des Tangors 'Temple' Minneola' (C. Jacquemond-INRA, 2012)	20
Figure 20 : Fruit de l'espèce <i>Citrus tangelo</i> 'Tangelo Minneola' (Dave's, 2015)	21
Figure 21 : Répartition mondiale de la production des agrumes par groupe (FAO, 2013).....	23
Figure 22 : Répartition des différents groupes d'agrumes selon leur production en Tunisie (GIF, 2016).....	24
Figure 23 : Carte de l'agrumiculture en Tunisie (Préparée par Amel Oueslati).....	25
Figure 24 : Assemblage des sccafoldes avec les marqueurs génétiques correspondants chez l'espèce <i>C. sinensis</i> (Xu <i>et al.</i> , 2012)	39

Figure 25 : Carte génétique circulaire du génome chloroplastique de l'espèce <i>C. sinensis</i> (Bausher <i>et al.</i> , 2006)	39
Figure 26 : Structure des sites des amorces désignées par Tabarlet <i>et al.</i> (1991) pour l'amplification de l'espaceur intergénique <i>trnL_(UAA)-trnF_(GAA)</i>	51
Figure 27 : Les étapes du génotypage par la méthode KASPar™ (Images de KBioscience, assemblage et montage par Amel Oueslati)	53
Figure 28 : Électrophorèse analytique sur gel d'agarose à 1% des ADNs génomiques totaux	65
Figure 29 : Profil d'amplification de l'espaceur intergénique <i>trnL-trnF</i>	67
Figure 30 : Exemple de visualisation d'une séquence de l'espaceur intergénique <i>trnL_{UAA}-trnF_{GAA}</i> du cultivar 'Orange Sanguine' de l'espèce <i>Citrus sinensis</i> par le logiciel BioEdit	68
Figure 31 : Vérification de l'identité des séquences <i>trnL_{UAA}-trnF_{GAA}</i> par comparaison avec les séquences publiées dans la banque de données GenBank de NCBI	68
Figure 32 : Un dendrogramme de type Neighbor-Joining des espèces tunisiennes du genre <i>Citrus</i> basé sur l'espaceur intergénique <i>trnL-trnF</i> . Les valeurs des Bootstrap (le pourcentage de bootstrap, >50%) résultant de 10 000 réplifications sont indiquées sur chaque branche ...	72
Figure 33 : La séquence nucléotidique (A) et la structure secondaire (B) du gène <i>trnF_(GAA)</i> (Pczai & Hyvönen, 2011)	76
Figure 34 : Alignement manuel des pseudogènes des espèces tunisiennes du genre <i>Citrus</i> démontrant leur structure et le polymorphisme observé entre les différentes espèces	77
Figure 35 : Spectre des fréquences des sites au niveau des séquences chloroplastiques de l'espaceur intergénique <i>trnL-trnF</i> chez les <i>Citrus</i> tunisiens	78
Figure 36 : Distribution des Mismatches des séquences chloroplastiques de l'espaceur intergénique <i>trnL-trnF</i> chez les <i>Citrus</i> tunisiens	79
Figure 37 : Distribution du nombre des SNPs diagnostiques au niveau des 54 clades considérés de Bayer <i>et al.</i> (2009)	83
Figure 38 : Arbre consensus basé sur les régions chloroplastiques montrant les liens de parenté au sein de la sous famille des Aurantioideae et des taxons outgroup (Bayer <i>et al.</i> , 2009)	84
Figure 39 : Un exemple de l'analyse KASPar du marqueur CP_49301 (A/G) distinguant les variétés du clade Q de Bayer <i>et al.</i> (2009)	89
Figure 40 : Classification des variétés de la sous famille des Aurantioideae en utilisant 27 marqueurs SNPs diagnostiques des clades Bayer <i>et al.</i> (2009)	92
Figure 41 : Procédure d'identification des marqueurs diagnostiques de la différenciation de <i>C. reticulata</i> / <i>C. maxima</i> à partir des données GBS	99
Figure 42 : Arbre issu de l'analyse NJ des variétés représentatives des mandarines et des pamplemousses	103

Figure 43 : Identification des introgressions interspécifiques chez les variétés représentatives des taxons ancestraux: exemple du chromosome 2 de la variété 'Ponkan'	104
Figure 44 : Comparaison de la distribution du G_{ST} entre les groupes de <i>C. reticulata</i> et de <i>C. maxima</i> basée sur les données initiales ($G_{ST\ 1}$) et les données issues après l'élimination des régions introgressées ($G_{ST\ ret-max}$) de type SNP et de type InDel	105
Figure 45 : Comparaison de la distribution des polymorphismes révélés par l'analyse GBS (Bleu), des DPs (Rouge) et des gènes (Vert) le long des neuf chromosomes.	106
Figure 46 : Analyse factorielle élaborée à partir du tableau de la dissimilarité des 54 variétés sur 30943 marqueurs	108
Figure 47 : Proportion globale de <i>C. maxima</i> et de <i>C. reticulata</i> chez les 55 variétés étudiées estimée à partir des polymorphismes diagnostiques	109
Figure 48 : Dédution de la structure phylogénomique du chromosome 2 de l'orange douce 'Washington Navel'	110
Figure 49 : Caryotype phylogénomique représentatif des neuf chromosomes	111
Figure 50 : Caryotype phylogénomique des neuf chromosomes des 6 variétés représentatives du taxon de base <i>C. maxima</i>	112
Figure 51 : Caryotype phylogénomique des neuf chromosomes des 11 variétés représentatives du taxon de base <i>C. reticulata</i>	113
Figure 52 : Caryotype phylogénomique des neuf chromosomes des mandarines hybrides	115
Figure 53 : Caryotype phylogénomique des neuf chromosomes des 7 tangors	116
Figure 54 : Caryotype phylogénomique des neuf chromosomes des 12 tangelos	118
Figure 55 : Caryotype phylogénomique des neuf chromosomes des tangelos hybrides	119
Figure 56 : Caryotype phylogénomique des neuf chromosomes des pomelos et des orangelos	120

Liste des Tableaux

Tableau 1 : Liste des variétés du genre <i>Citrus</i> classées selon la classification de Swingle & Reece (1967)	45
Tableau 2 : Liste des espèces et des accessions de la sous famille des Aurantioideae utilisées dans l'étude <i>in silico</i> et KAPar expérimentale	47
Tableau 3 : Liste des espèces et des accessions du groupe des agrumes vrais utilisées dans l'étude <i>in silico</i> et KAPar expérimentale	48
Tableau 4 : Liste des espèces et des variétés du genre <i>Citrus</i> analysées	50
Tableau 5 : Caractéristiques des séquences des espèces du genre <i>Citrus</i>	69
Tableau 6 : Estimation du maximum de vraisemblance du modèle de la substitution nucléotidique au niveau de l'espaceur intergénique <i>trnL-trnF</i>	70
Tableau 7 : Paramètres du polymorphisme testés sur les séquences de l'espaceur intergénique <i>trnL-trnF</i>	71
Tableau 8 : Tests de neutralité sélective au niveau des séquences chloroplastiques de l'espaceur intergénique <i>trnL-trnF</i> chez les espèces tunisiennes du genre <i>Citrus</i>	78
Tableau 9 : Nombre des SNP diagnostiques des clades de Bayer <i>et al.</i> (2009).....	87
Tableau 10 : Résultats des analyses des marqueurs KASPar	90
Tableau 11 : Polymorphisme issue des données GBS des 55 variétés étudiées	101
Tableau 12 : Polymorphismes (SNP et InDel) identifiés à partir des données GBS chez les 11 variétés de mandarines et les 6 variétés de pamplemousses	102
Tableau 13 : Distribution des DPs entre les neuf chromosomes	106

Liste des Annexes

Annexe 1 : Liste des séquences des accessions utilisées dans l'analyse <i>in silico</i> des SNP	157
Annexe 2 : Liste du matériel végétal utilisé dans l'analyse KASPar	161
Annexe 3 : Liste des SNP sélectionnés et leurs séquences avec 50 pb de chaque côté du SNP.....	165
Annexe 4 : Exemple d'un SNP diagnostique localisé au niveau de la région <i>rps4-trnT</i> à la position 49301 distinguant les accessions du clade Q (A) de toutes les autres accessions de la sous famille des Aurantioideae (G)	167
Annexe 5 : Liste du matériel végétal utilisé	168
Annexe 6 : Les régions d'introgressions chez les variétés représentatives des taxons de référence <i>Citrus reticulata</i> et <i>Citrus maxima</i>	170
Annexe 7 : Analyse de l'hétérozygotie observée et estimation des données manquantes chez les variétés étudiées au niveau de chaque chromosome	172
Annexe 8 : Analyse de l'homozygotie et de l'hétérozygotie des taxons de références chez les variétés étudiées à partir des points diagnostiques de chaque chromosome	174
Annexe 9 : Analyse de l'hétérozygotie des taxons de références chez les variétés étudiées à partir des points diagnostiques de chaque chromosome	176

Introduction Générale

*« La science consiste à passer d'un étonnement à un autre »
Aristote*

Le secteur agricole revêt une place prépondérante dans le tissu économique mondial. L'agrumiculture, située dans les régions à climats tropicaux et subtropicaux ainsi que dans les zones s'étendant de 40° au Nord à 40° au Sud de l'équateur, est considérée comme la première production fruitière au monde avec plus de 130 millions de tonnes. La Tunisie qui produit 431 milles tonnes/an, marque des produits associés à son image comme les orangers 'Maltaise', mondialement connus et unanimement appréciés pour leur saveur unique.

Les agrumes sont, en effet, des arbres fruitiers originaires de l'Asie. Ces arbres sont réputés pour leur très beau feuillage vert et leurs fleurs blanches et parfumées qui en font un très bel arbre ornemental.

Le groupe des agrumes dit vrais appartient à la sous famille des Aurantioideae et compte six genres interféquents: *Poncirus* Raf., *Fortunella* Swingle, *Microcitrus* Swingle, *Clymenia* Swingle, *Eremocitrus* Swingle et *Citrus* (L.). Le genre *Poncirus* est généralement utilisé comme porte-greffe et le genre *Citrus* comprend les fruits comestibles. Ainsi, ce dernier comprend quatre espèces principales *Citrus reticulata* Blanco (mandarine), *Citrus maxima* (Burm.) Merr (pamplemousse), *Citrus medica* L. (cédrat) et *Citrus micrantha* Wester (papeda). Les autres espèces (orange douce, orange amère, clémentine, lime, citron, pomelo) sont considérées comme des espèces secondaires originaires d'hybridations interspécifiques entre les quatre taxons de base.

Bien que les agrumes représentent une biodiversité et une production importante à l'échelle mondiale et nationale avec une large diversité du sol et du climat, cette culture est confrontée à différentes contraintes biotiques et abiotiques selon les régions de production. En effet, au niveau du Bassin Méditerranéen, la culture des agrumes est touchée par la salinité et la sécheresse. Ils sont les deux principaux facteurs abiotiques limitant du développement agrumicole qui entraînent un déficit hydrique et la chute des feuilles. De même, la nature du sol peut affecter cette production qui pourrait être la cause de maladies touchant les agrumes comme le cas de la Chlorose ferrique. D'autres parts, l'agrumiculture est affectée par différents stressés biotiques qui causent une réduction de la production et de la qualité du fruit, qui peuvent même entraîner une destruction de l'industrie agrumicole. Les maladies présentes dans la majorité des régions sont causées par les nématodes, l'oomycète *Phytophthora* sp. ou le virus de la Tristeza (CTV : *Citrus Tristeza Virus*) (Ollitrault & Navarro, 2012). La Tunisie est confrontée aussi à la maladie de brindilles sèches (Malsecco) ce qui empêche d'une part l'utilisation de certains porte-greffes et d'autre part, réduit la production et la qualité des variétés spécifiques (GIF, 2016).

Etant donné, la richesse agrumicole nationale et internationale avec les importantes contraintes qui touchent cette culture, des efforts importants ont été consentit pour

développer ce patrimoine fruitier. Une stratégie de développement et d'amélioration bénéfique au secteur des agrumes a été mise en place. L'analyse du germoplasme local des porte-greffes d'agrumes était une étape préliminaire avant de mettre en œuvre les programmes d'amélioration des agrumes dans le but d'identifier les porte-greffes adaptables à chaque région du pays (Snoussi *et al.*, 2006). Cette analyse a été complétée par un programme d'amélioration variétale, objectif majeur de l'agrumiculture. Ainsi, la Tunisie a suivis un assainissement des variétés anciennes et une introduction de nouvelles variétés afin de viser des dates de maturité échelonnées ce qui permet d'enrichir les étalages des variétés tardives et donc étaler la période de production offrant une production variée et de bonne qualité. Au niveau du bassin méditerranéen, l'amélioration variétale par la voie de recombinaison sexuée était la principale stratégie attribuée. Durant les dernières années, l'amélioration variétale s'est intéressée à des critères de sélection tel que: l'absence de pépin, la stérilité mâle et femelle, l'auto-incompatibilité et la parthénocarpie (Ollitrault & Navarro, 2012).

Cette stratégie d'amélioration variétale était limitée à cause de certains caractères biologiques des agrumes. En effet, l'apomixie a engendré une accumulation d'allèles récessifs néfastes ou délétères ainsi que la formation de pépins polyembryonnés limitant souvent la formation de l'embryon zygotique. La stérilité a été la cause de l'aspermie de cultivars sélectionnés. Ainsi, l'incompatibilité sexuelle, les structures génomiques d'origine interspécifique hautement hétérozygote, l'apomixie et la stérilité entraînent, chez les agrumes, la perte du morphotype après recombinaison sexuée. En outre, l'évolution réticulée pose un important défi dans l'étude de la phylogénie. Cependant, la spéciation, l'hybridation, l'introgression et le transfert latéral de gènes contribuent à une histoire évolutive qui ne peut pas être représentée sous une forme d'arbre phylogénétique. En plus, quand la propagation végétative prend la place directement ou après quelques générations d'un événement de réticulation, plusieurs parties du génome restent souvent en hétérozygotie interspécifique. Par ailleurs, l'apomixie partielle de la majorité des espèces secondaires des *Citrus* a été un élément essentiel limitant la recombinaison interspécifique. Cependant, la majorité des *Citrus* cultivés ont des structures génomiques en mosaïque interspécifiques hautement hétérozygotes issues de larges fragments chromosomiques des génomes de taxons de base (Grosser & Gmitter, 1990). Ceci montre que les agrumes pourraient être un bon exemple de pool génétique avec une évolution réticulée.

Afin de contribuer à lutter contre ces contraintes génétiques et approfondir ces travaux pour améliorer et protéger cette ressource génétique vue sa place prépondérante dans l'économie nationale et internationale, l'amélioration variétale par la voie d'amélioration conventionnelle, la recombinaison sexuée, n'est plus envisageable. Au bassin

mediterranéen, de nouvelles stratégies d'amélioration variétale ont été mises en place telles que : la ploïdie, le recours aux cybrides, la sélection de mutants spontanés, la mutagenèse induite/dirigée, la transformation génétique et la transgénèse.

Une analyse phylogénétique du germoplasme d'agrumes devient une nécessité pour le sauvegarder. En effet, le recours aux analyses moléculaires dans les études de la diversité génétique des phytoressources permet de mieux identifier et caractériser ces ressources pour mettre en œuvre des acteurs permettant de les améliorer et augmenter ainsi la productivité agricole, dresser des stratégies d'échantillonnage, évaluer l'importance de cette diversité génétique et donc renforcer les études visées au développement du secteur agricole. Une bonne connaissance de l'origine génétique des agrumes ouvre la voie vers une étude génomique permettant de développer une stratégie d'amélioration. Celle-ci vise à reconstruire les structures génomiques des variétés secondaires hybrides à partir des pools génétiques des taxons de base qui sont le réservoir de diversité des formes cultivées.

Dans le cadre des travaux présentés dans cette thèse, nous nous sommes intéressés à plusieurs approches d'analyse pour répondre à ces objectifs en apportant de nouvelles connaissances dans, l'organisation génétique et l'origine phylogénétique du genre *Citrus*, allant aux genres proches, au groupe des agrumes vrais jusqu'à atteindre la sous-famille des Aurantioideae et dans le décryptage des structurations phylogénomiques du germoplasme des variétés modernes de mandariniers et d'hybrides de mandariniers.

Au niveau du premier chapitre, une étude du polymorphisme moléculaire chez sept espèces tunisiennes du genre *Citrus* a été menée par l'analyse des séquences de l'espaceur intergénique chloroplastique *trnL_{UAA}-trnF_{GAA}*. Afin d'établir leur lien phylogénétique et leur origine maternelle en essayant de chercher à situer le germoplasme tunisien dans une diversité plus globale représentée par des variétés de références et de tracer leur histoire évolutive. En effet, l'identification et la caractérisation des variétés d'agrumes tunisiennes nous permet de mettre en place une collection de référence et de préparer une base génétique large. Ceux-ci permettra de faciliter les travaux d'amélioration et de création de nouvelles variétés en vue de satisfaire tant le marché local que celui de l'export et garantir la pérennité des espèces.

Au niveau du deuxième chapitre, la présente thèse envisage à focaliser l'analyse moléculaire dans un polymorphisme ciblé avec une importante valeur taxonomique. Notre hypothèse repose sur le fait qu'avec une conservation élevée de la séquence chloroplastique, l'identification de polymorphisme de mutation ponctuelle (SNP) spécifique des clades devrait permettre de fournir une clé moléculaire taxonomique au niveau de la sous famille des Aurantioideae en se basant sur une simple, robuste et rentable méthode de génotypage des marqueurs SNPs. Cet ensemble de marqueurs diagnostiques SNPs doit être par la suite

appliqué à une large collection de germoplasmes. Pour ce faire, une analyse *in silico* de huit régions chloroplastiques de 79 accèsions de la sous famille des Aurantioideae a été effectuée pour identifier et sélectionner les SNPs diagnostiques des différenciations taxonomiques aux niveaux inter-tribu, inter-sous tribu, inter-genre et interspécifiques. L'analyse de ce polymorphisme avec la méthode de génotypage permettant une transférabilité fluide au sein de la sous famille des Aurantioideae a été envisagée. Ces marqueurs ont été testés via la technique PCR compétitive allèle spécifique (la technologie KASPar), qui est considérée comme une méthode efficace pour identifier les polymorphismes de type SNP. Cette étude novatrice, a permis de développer, pour la première fois au monde des plantes, des marqueurs KASPAar développés à partir de marqueurs SNPs chloroplastiques sélectionnés pour valider leur valeur taxonomique, leur transférabilité à l'ensemble de la sous famille des Aurantioideae et de définir la phylogénie maternelle de différentes espèces du genre *Citrus*.

Cependant, la phylogénie chloroplastique ne révèle qu'une part de l'histoire évolutive du pool génétique due à son hérédité monoparentale. Pour les groupes très récemment divergés, l'information génétique chloroplastique n'est pas suffisamment résolutive à cause de leur faible taux de variation. Pour cette raison, les données chloroplastiques et nucléaires sont complémentaires dans la clarification de la structure phylogénomique et de l'histoire évolutive des groupes récemment divergés particulièrement quand les événements de réticulation sont fréquents tel que le cas des *Citrus*. Le recours au séquençage de nouvelle génération couplé à la réduction de la complexité du génome en utilisant la méthode du Génotypage par Séquençage (GBS) serait convenable pour définir la phylogénomie nucléaire complémentaire des SNPs diagnostiques chloroplastiques.

Nous nous sommes, ainsi, passés dans la présente thèse de la phylogénie à la phylogénomie. En effet, le troisième chapitre de cette thèse, propose une analyse de la structure génomique des variétés des petits agrumes comprenant les mandariniers hybrides, les tangors et les tangelos (et les tangelos hybrides). Pour ce faire, une analyse par la technique GBS en utilisant le séquençage *Illumina run hiseq 2000* a été effectuée. Une identification des polymorphismes diagnostiques de la différenciation des variétés représentatives des taxons ancestraux *C. maxima* et *C. reticulata* a été élaborée et ont été par la suite utilisés dans l'analyse de la phylogénomie des segments chromosomiques de l'ensemble du génome.

Etude Bibliographique

1. Origine géographique et histoire des Agrumes

Le centre d'origine des agrumes est reconnu unanimement très vaste, ainsi, il inclut le Sud de la Chine, le Nord-Est de l'Inde, la Péninsule Indo-Chinoise et la Birmanie, mais la localisation du premier centre d'origine reste controversée. Ce centre d'origine varie, selon les chercheurs, de la région montagneuse du Sud de la Chine et du Nord-Est de l'Inde (Tolkowsky, 1938) jusqu'au Nord-Est de l'Inde et Burma (Tanaka, 1954) ou jusqu'à la Province de Yunnan de la Chine (Gmitter & Hu, 1990).

La culture des espèces du genre *Citrus* remonte à plusieurs milliers d'années dans le Sud-Est de l'Asie puis a connu une diffusion dans le monde (Figure 1). La culture des agrumes a connu une part de l'histoire ancienne de la Chine et des histoires sur la mandarine et le pamplemousse ont été mentionnées dans le chapitre "Tribute to Yu" (2205–2197 avant Jésus-Christ (J.C)). De même, la collection religieuse de l'Inde "Vajaseneyi sambita" écrite en 800 ans avant J.C, a abordé cette culture en décrivant les cédrats et les citrons. Les cédrats sont les premières espèces reconnues dans le Bassin Méditerranéen, probablement introduits de la Perse et de la Grèce à partir de l'Inde par Alexandre le Grand aux environs 300 ans avant J.C. Les mosaïques antiques témoignent la connaissance des citrons par les Romains aux environs 100 ans avant J.C sans avoir de preuves de culture de ce fruit. Lors de l'expansion de l'Empire Arabe, dans le dixième et l'onzième siècle, les arabes ont diffusé les agrumes dans l'Europe et l'Afrique du Nord. Les cédrats, les oranges amères, les citrons, les limes et les pamplemousses ont été décrits dans les livres du X^e et XI^e siècle de l'Espagne. Au début du XV^e siècle les premières cultures de l'orange douce ont été élaborées en Europe, probablement introduite par les Génois, mais elles n'ont pas été diffusées qu'à la sélection de nouvelles espèces par les Portugais au début du XVI^e siècle. Alors que les mandarines cultivées au Sud-Est Asiatique n'ont été introduites en Europe qu'au XIX^e siècle (Ollitrault & Navarro, 2012).

Au Japon, la présence des agrumes date du premier siècle et même, certaines espèces ont été introduites en Chine et d'autres pays voisins par une expédition japonaise. Les agrumes ont été introduits en Amérique par Christopher Columbus dans son deuxième voyage en 1483. Il a pris des graines des oranges, des citrons et des cédrats à l'île Hispaniola et à partir de laquelle les fruits ont été diffusés vers les autres îles de l'Amérique Centrale. Entre 1515 et 1565, des explorateurs espagnols ont introduits les agrumes à la Floride et approximativement durant la même période, les Portugais ont introduits les agrumes au Brésil. En 1654, les Portugais les ont introduits à l'Afrique de l'Ouest et la colonie Hollandaise a introduit les oranges douces à l'Afrique du Sud.

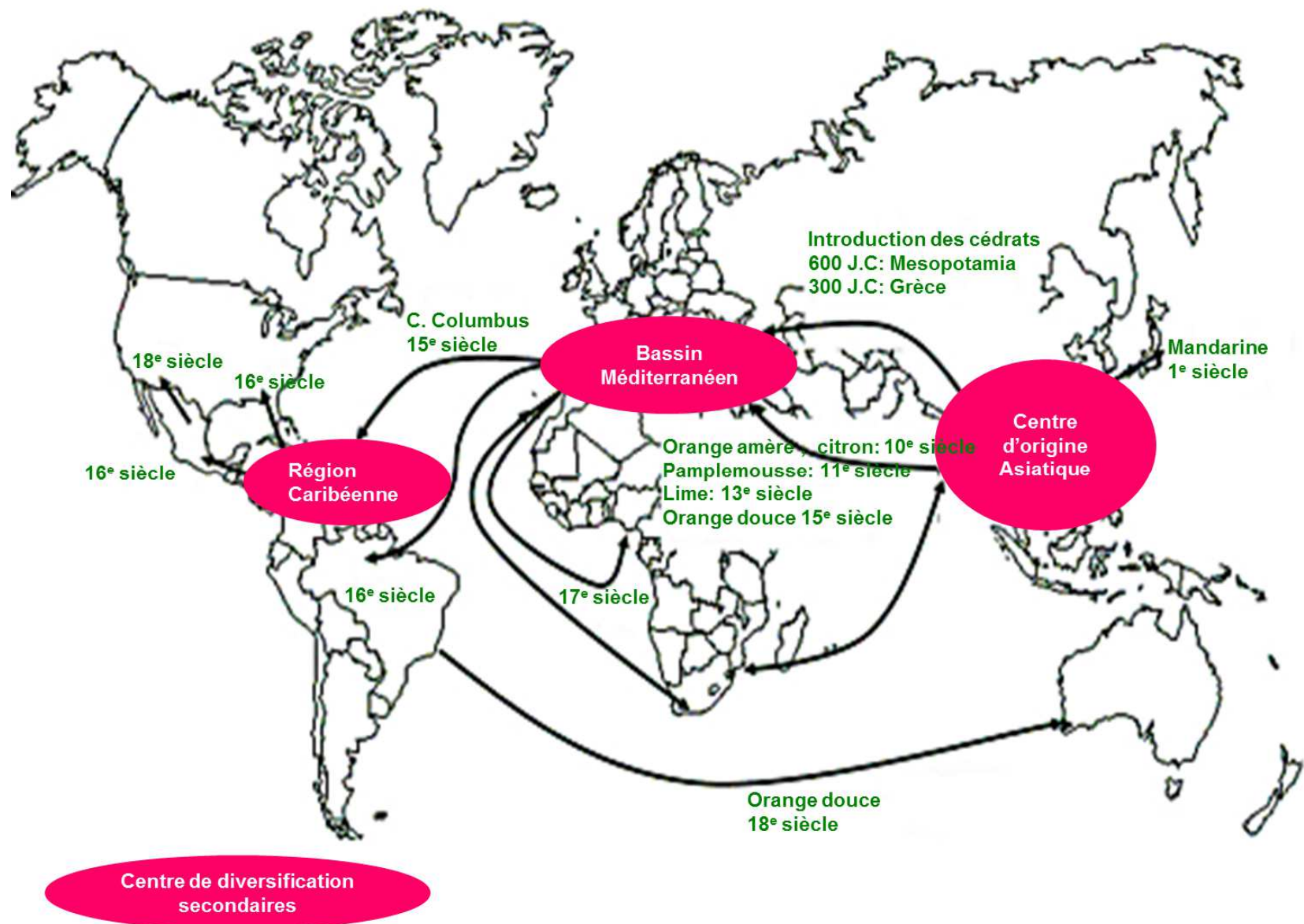


Figure 1 : Carte représentative de l'origine et de la diffusion des agrumes dans le monde (Praloran, 1971)

En 1769, les colons de la première flotte ont apporté des oranges, des citrons verts, et les citrons du Brésil à l'Australie. Les oranges et les citrons ont été introduits en Californie en 1769 (Ollitrault & Navarro, 2012).

L'orange douce est un fruit mondialement consommé (Xu *et al.*, 2013). Il a reconnu une importante diversification grâce à des mutations. Il a été introduit relativement tard dans le bassin Méditerranéen (Aubert, 2001). Sa diversification a reconnue trois zones importantes. La principale zone de diversification est la Péninsule Ibérienne caractérisée par l'orange 'Cadenera'. Le Portugal et l'Espagne sont caractérisés par la culture de l'orange 'Valencia late' et l'orange 'Navel'. La seconde zone de diversification des oranges douce couvre, la Tunisie, la Malta et la Sicile d'où l'origine de l'orange 'Sanguine' et 'Demi-Sanguine'. L'orange 'Tarocco' est la plus cultivée dans cette zone suivie par la 'Maltaise' largement cultivée en Tunisie comme le cultivar tardé 'Barlerin'. Cette culture fortement répandue dans la péninsule du Cap Bon, terre réputée fertile depuis l'Antiquité, a attiré les jardiniers Andalous qui ont y cultivé la fameuse orange 'Maltaise' et 'Demi Sanguine' de la Tunisie. Le troisième centre de diversification est le proche orient caractérisé par l'orange 'Blond' (Ollitrault & Navarro, 2012).

2. Etude Botanique et description morphologique

2.1 La famille des Aurantioideae

La famille des Rutaceae compte 158 genres et 1900 espèces (Mabberley, 2008). Elle regroupe sept sous familles dont fait partie les Aurantioideae (Engler, 1931) considérés comme la plus riche sous famille (Thorne, 2000). Selon la classification de Swingle & Reece (1967) largement adoptée par la communauté scientifique de la recherche des Agrumes, les Aurantioideae regroupent 33 genres et 210 espèces. Cette sous famille est subdivisée en deux tribus : les Clauseneae qui forment cinq genres et les Citreae 28 genres (Figure 2).

Swingle & Reece (1967) pensent que la tribu des Clauseneae contient les genres les plus primitifs des Aurantioideae. En se basant sur leur classification, la tribu des Clauseneae regroupe trois sous-tribus : les Micromelineae (un seul genre), les Clauseninae (trois genres) et les Merrillinae (un seul genre). Ainsi, toutes les espèces appartenant à la tribu des Clauseneae n'ont pas d'épine axillaire et possèdent des feuilles imparipennées avec une foliole alternativement attachée au rachis. Les fruits, sauf chez le genre *Merrilla*, sont généralement petits, demi-secs ou aux baies juteuses. *Merrilla* est le seul genre de la sous famille des Aurantioideae possédant des fleurs zygomorphiques (Figure 2).

Quant à la tribu des Citreae, elle regroupe : les Triphasilineae (huit genres), les Balsamocitrinae (sept genres) et les Citrinae. Ainsi, toutes les espèces de cette tribu développent des épines axillaires, individualisées ou jumelées et parfois courbées comme

chez les genres *Luvunga* et *Paramignya*. La simplicité de leurs feuilles typiques unifoliées ou trifoliées les distingue des Clauseneae. Mais, quelques genres (*Feronia*, et *Hesperthusa*) de la tribu des Citreae présentent des feuilles imparipennées avec une foliole inversée. La sous-tribu des Citrinae est considérée la plus importante ; elle est distinguée des autres sous-tribus par la présence de vésicules de pulpes dans leurs fruits. Cette sous-tribu contient trois groupes : les genres à fruits primitifs (cinq genres), les genres à fruits proches des agrumes (deux genres) et les genres des agrumes vrais (six genres). Ce dernier groupe inclut l'ensemble des formes cultivées et compte le genre *Citrus* L. et les genres qu'y sont reliés : *Fortunella* Swingle, *Microcitrus* Swingle, *Eremocitrus* Swingle, *Clymenia* Swingle and *Poncirus* Raf. (Swingle & Reece, 1967 ; Krueger & Navarro, 2007 ; Morton *et al.*, 2008) (Figure 2).

Ainsi, la classification des tribus et des sous-tribus de la sous famille des Aurantioideae a été un débat critique. En effet, Tanaka (1936) a subdivisé la sous famille en huit tribus et huit sous-tribus formant 28 genres alors qu'Engler (1931) a regroupé l'ensemble de la sous famille en une seule tribu appelée Aurantieae subdivisée en deux sous-tribus : Hesperethusinae (16 genres) et Citrinae (13 genres). Récemment, Mabberley (1997) a fusionné *Eremocitrus*, *Fortunella*, et *Microcitrus* avec *Citrus* et a suggéré que *Poncirus* doit rester un genre à part. Cependant, ces différentes classifications ont été basées sur des caractères morphologiques. Dans cette dernière décennie, l'outil moléculaire était performant dans la clarification des liens de parenté de la sous famille des Aurantioideae.

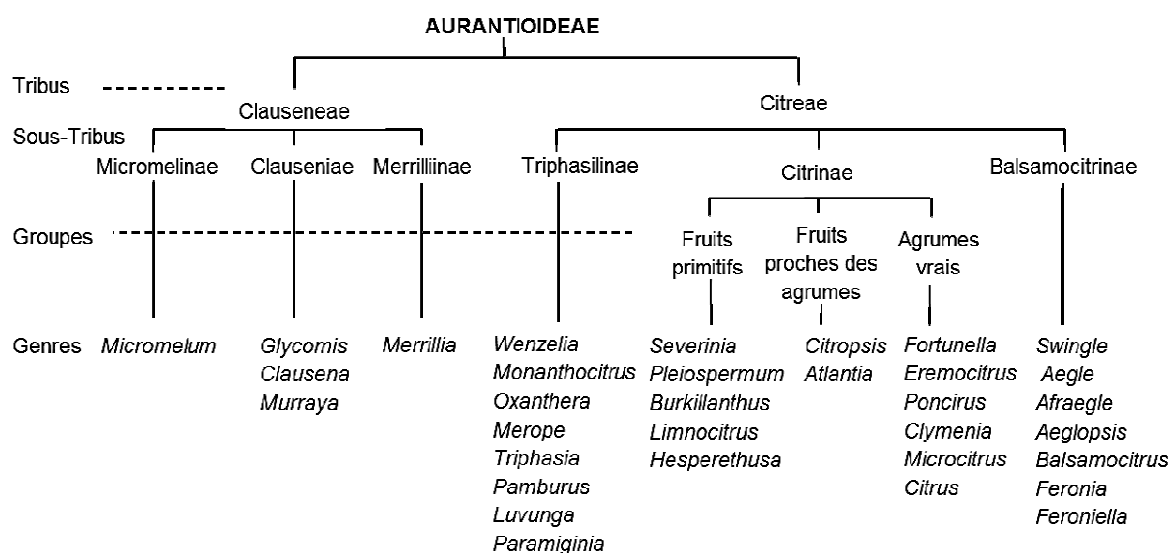


Figure 2 : Taxonomie de la sous famille des Aurantioideae (Swingle & Reece, 1967)

2.2 Le groupe des Agrumes vrais

Le groupe des agrumes vrais est formé par six genres parmi lesquels trois jouent un rôle central en horticulture comme variété ou porte-greffe : *Citrus*, *Fortunella* et *Poncirus*.

➤ Le genre *Fortunella* Swingle

Le genre *Fortunella* désigne les kumquats. Il regroupe quatre espèces : *Fortunella japonica* (Thunb.) Swing., *Fortunella hindsii* (Champ.) Swing., *Fortunella polyandra* (Ridl.) Tan., et *Fortunella margarita* (Lour.) Swing., (Figure 3). Il s'agit de petits fruits à écorce comestible douce possédant de trois à sept ovaires loculaires et des vésicules minces. Les feuilles sont rugueuses riches en glandes et en huiles essentielles. L'arbre du kumquat est à feuilles persistantes; avec une taille variable. Cet arbre est tolérant au froid, il présente une floraison tardive par rapport aux autres espèces d'agrumes. Les kumquats sont aussi à utilisation ornementale.



Figure 3 : Fruit de l'espèce *Fortunella margarita* 'Nagami kumquat' (Aujardin, 2016)

➤ Le genre *Poncirus* Raf.

Le genre *Poncirus* est formé par une seule espèce *Poncirus trifoliata* (L.) Raf. (Figure 4). Les *Poncirus* se distinguent des autres agrumes par leurs feuilles caduques et trifoliolées. Ils présentent la floraison la plus précoce des agrumes et possèdent six à huit ovaires loculaires. Ils produisent des fruits non comestibles avec des vésicules très riches en huiles essentielles, et jouent un rôle important en agrumiculture comme porte-greffe (Swingle & Reece, 1967).



Figure 4 : Fruit de l'espèce *Poncirus trifoliata* (Thibault, 2016)

➤ Le genre *Eremocitrus* Swingle

Eremocitrus est un genre monospécifique il comprend une seule espèce *Eremocitrus glauca* (Lindl.) Swing (Figure 5). Les *Eremocitrus* sont des plantes à feuilles vertes-grises, épaisses et poilues sur les deux côtés. Les fleurs sont simples ou en groupes, avec trois à cinq ovaires loculaires. Les fruits sont des vésicules minces ovoïdes ou piriformes. Ils tolèrent le froid et la sécheresse (Swingle & Reece, 1967).



Figure 5 : Fruit de l'espèce *Eremocitrus glauca* (Lucas, 2015)

➤ Le genre *Microcitrus* Swingle

Le genre des *Microcitrus* est constitué de six espèces dont cinq originaires de l'Est de l'Australie : *Microcitrus australis* (Planch.) Swing., *Microcitrus garrowayi* (F.M. Bail.) Swing., *Microcitrus maideniana* (Domin) Swing., *Microcitrus inodora* (F. M. Bail.) Swing. et *Microcitrus australasica* (F. Muell.) Swing. (Figure 6) quant à la *Microcitrus warburgiana* (F. M. Bail.) Tan. est originaire du Sud-Est de la Nouvelle Guinée. Les *Micocitrus* sont à feuillage dimorphe, avec de petites fleurs. Les arbres sont sous forme d'arbuste, les fruits sont généralement allongés. Ils sont semi-xérophiles et peuvent résister à de longues sécheresses (Swingle & Reece, 1967).

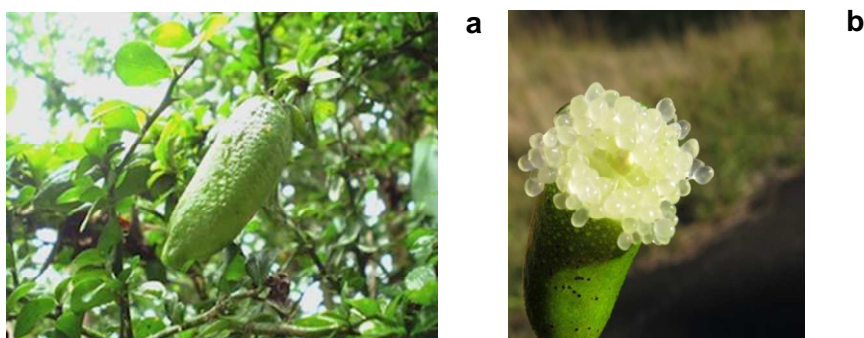


Figure 6 : Fruit de l'espèce *Microcitrus australisica* 'Citron caviar' (a), Vésicules d'un fruit 'Citron caviar' coupé en deux (b) (Clémentine, 2012)

➤ Le genre *Clymenia Swingle*

Clymenia plus probablement le genre le plus primitif au sein des agrumes vrais ne comprend qu'une seule espèce *Clymenia polyandra* (Tan.) Swing. Elle a des feuilles qui ressemblent à ceux des *Triphasinae*. Les fleurs sont de forme allongée formées de dix à vingt étamines et sépales (Swingle & Reece, 1967). Le fruit est représenté au niveau de la figure 7.

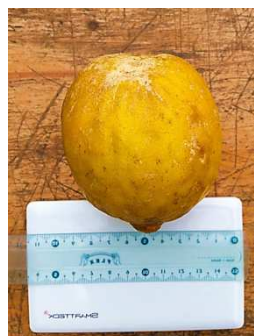


Figure 7 : Fruit de l'espèce *Clymenia polyandra* (Wakonai, 2016)

➤ Le genre *Citrus L.*

Le genre *Citrus* comprend la plupart des agrumes cultivés. Ils se caractérisent par leurs fruits et leurs huiles essentielles. Les fruits des principales espèces et variétés cultivées des *Citrus* diffèrent par leur coloration, leur forme, leur calibre, la composition de leur jus et leur période de maturité (Swingle & Reece, 1967).

2.3 Le genre *Citrus L.*

La plupart des agrumes cultivés font partie du genre *Citrus* qui, selon les taxonomistes, regroupe entre 16 (Swingle & Reece, 1967) et 156 espèces (Tanaka, 1961). La classification de Swingle & Reece (1967), basée sur la comestibilité des fruits, distingue entre le sous genre des *Eucitrus*, où tous les taxons cultivés sont regroupés, et le sous genre des *Papeda*.

Le sous genre dernièrement cité est constitué de six espèces : *Citrus micrantha* Wester, *Citrus ichangensis* Swing. , *Citrus hystrix* (D.C.), *Citrus latipes* Swing. (Tan.), *Citrus celebica* Koord et *Citrus macroptera* Montr. (Satkara). Le sous genre *Eucitrus* comprend dix espèces cultivées: *Citrus medica* L. (Cédrat), *Citrus aurantium* L. (Bigarade), *Citrus sinensis* (L.) Osb. (Orange douce), *Citrus limon* (L.) Burm. (Citron), *Citrus aurantifolia* (Christm) Swing. (Lime), *Citrus maxima* (Burm.) Merr. (Pamplemousse), *Citrus paradisi* Macf. (Pomelo), *Citrus reticulata* Blanco (Mandarine) *Citrus indica* Tan. et *Citrus tachibana* (Mac.) Tan. Les deux dernières espèces sont moins importantes sur le plan horticole.

La taxonomie de Tanaka (1961) est beaucoup plus détaillée que celle adoptée par Swingle & Reece (1967). En effet, Tanaka a subdivisé le genre *Citrus* en deux sous genres :

Archicitrus et *Metacitrus*. Ainsi, les principales différences entre la classification de Swingle & Reece et celle de Tanaka concernent les citrons, les limes et les mandarines.

Cette classification relativement complexe résulte de l'association d'une large diversité morphologique, d'une compatibilité sexuelle interspécifique totale au sein du genre et de l'apomixie. Cette forme d'apomixie, fixant les structures génétiques complexes à travers la propagation de la graine, a conduit certains taxonomistes à considérer les familles clonales d'origine interspécifique comme de nouvelles espèces.

2.3.1 Origine génétique des taxons de base

Scora (1975) et Barrett & Rhods (1976), ont proposé l'existence de trois espèces de base du genre *Citrus* en se basant sur des descripteurs morphologiques: *Citrus maxima* (le pamplemoussier), *Citrus medica* (le cédratier) et *Citrus reticulata* (le mandarinier) à partir desquelles provenaient tous les agrumes cultivés. La forte organisation de la variabilité phénotypique autours de ces trois taxons de référence a été en outre confirmée par des caractères morphologiques végétatifs (Ollitrault *et al.*, 2003) et la composition en caroténoïdes (Fanciullino *et al.*, 2006). Des études protéiques (Handa *et al.* 1986), isoenzymatiques (Herrero *et al.* 1996, 1997 ; Ollitrault *et al.*, 2003) et moléculaires: RFLPs, RAPDs (Federici *et al.*, 1998 ; Nicolosi *et al.*, 2000 ; Fanciullino *et al.*, 2007) , SSRs (Luro *et al.*, 2001 ; Barkley *et al.*, 2006) et plastomiques (Green *et al.*, 1986) ont validé cette hypothèse.

La différenciation entre ces taxons interféconds et isolés géographiquement pourrait être expliquée par la spéciation allopatrique par effet fondateur. Le pamplemousse originaire de l'Archipel Malais et de l'Indonésie, le cédrat évolué au Nord-Est de l'Inde et les régions voisines de la Birmanie et de la Chine quant à la mandarine diversifiée dans les régions de la Vietnam, le Sud de la Chine et au Japon (Scora, 1975 ; Webber, 1967). Des études postérieures (Federici *et al.*, 1998; Nicolosi *et al.*, 2000 ; Ollitrault *et al.*, 2012) ont montré la contribution de *Citrus micrantha* (la papéda) dans l'apparition de nouvelles variétés du genre *Citrus* spécifiquement le citron vert et l'ont considéré ainsi comme le quatrième taxon de référence. Par ailleurs, toutes les autres espèces cultivées apparaissent par une recombinaison entre les quatre taxons de base.

➤ Les pamplemoussiers

Le pamplemousse (*C. maxima*, Figure 8) est un fruit diversifié, il pourrait être pigmenté ou non pigmenté, acide ou non acide. Il est plutôt monoembryonique qu'apomictique. Le pamplemousse est principalement à pollinisation-croisée mais la présence de certaines formes produisant une descendance rigoureuse relativement autoféconde n'est pas exclue (Moore *et al.*, 2001). D'autre part, selon Barrett & Rhodes

(1976), une affinité intra-spécifique assez élevée a été enregistrée chez les pamplemoussiers ainsi que la propension de l'autofécondation qui sont due à l'homozygotie de l'espèce.



Figure 8 : Fruit de l'espèce *Citrus maxima* (Berger, 2003)

➤ Les cédratiers

Les cédrats (*C. medica*) sont classés en variétés acides et douces principalement cultivées pour leurs zestes. La figure 9 représente la variété 'Main de Bouddha' une des principales variétés des cédrats. Historiquement, le fruit a une signification religieuse importante pour certaines cultures. Le cédrat est monoembryonique et faiblement hétérozygote (Moore *et al.*, 2001).



Figure 9 : Fruit de l'espèce *Citrus medica* 'Main de Bouddha' (Ly, 2014)

➤ Les mandariniers

Les mandarines (*C. reticulata*, Figure 10) sont le groupe des *Citrus* le plus hétérogène phénotypiquement. Ainsi, il existe des variétés douces, à utilisation gastronomique, et des variétés amères, utilisées comme porte-greffe et comme produit aromatique et médicinal. D'autres part, elles regroupent des variétés monoembryoniques et polyembryoniques, tout de même des formes auto-fertiles et auto-incompatibles. Les mandarines présentent une hétérozygotie beaucoup plus complexe que les autres formes apomictiques (Barrett & Rhodes, 1976).



Figure 10 : Fruit de l'espèce *Citrus reticulata* (Amel Oueslati, 2010)

➤ Les papedas

La papeda (*C. micrantha*, Figure 11) comprend les espèces sauvages des agrumes. Elles sont à pétioles longs et ailés. Les fleurs sont petites, avec des étamines libres, car les faisceaux latéraux des sépales et la nervure médiane du pétale n'ont pas été fusionnés. Dans les fruits, les vésicules de pulpes ont de nombreuses gouttes d'huile acre, qui sont à l'origine du goût amer du fruit, ce qui rend les espèces des papedas non comestibles.



Figure 11 : Fruit de l'espèce *Citrus micrantha* (Robyn Jay, 2016)

2.3.2 Origine génétique des variétés secondaires

Les résultats obtenus avec les outils moléculaires ont conforté l'hypothèse de la dérivation des nouvelles variétés des hybridations entre les quatre espèces de base ou des mutations naturelles (Figure 12), et ont par ailleurs permis de clarifier les liens de parenté entre les espèces secondaires (Nicolosi *et al.*, 2000 ; Garcia *et al.*, 2012 ; Ollitrault *et al.*, 2012 ; Garcia *et al.*, 2013). Avec l'utilisation des marqueurs biochimiques comme les isoenzymes (Herrero *et al.*, 1996 ; Ollitrault *et al.*, 2003) et des marqueurs moléculaires codominants tels que les RFLP (Federici *et al.*, 1998; Fanciullino *et al.*, 2007) et les SSR (Luro *et al.*, 2003 ; Barkley *et al.*, 2006) la plupart des espèces secondaires font preuve d'un haut niveau d'hétérozygotie avec l'absence de polymorphisme inter-variétal au niveau interspécifique. Ce sont de « fausses espèces » typiques associées à la multiplication végétative par apomixie ou greffage, où la diversité phénotypique variétale est produite par une accumulation de variations à partir d'un prototype ancestral hybride sans recombinaison

sexuelle. Avec ces espèces il y a une discordance entre le phénotype et la diversité moléculaire. Bretó *et al.* (2001) ont proposé que les éléments transposables puissent être une importante source de variation des espèces du genre *Citrus* par l'outil de la propagation végétative.

L'étude actuelle de la diversité des formes cultivées montre la maintenance de la robustesse des structures au niveau moléculaires et morphologiques. Ceci indique que l'échange génétique entre les taxons de base a été limité. L'apomixie associée à la polyembryonie de la plupart des espèces secondaires est un élément essentiel dans la limitation des flux géniques. D'autres facteurs comme la différenciation structurale des génomes ont aussi favorisé la maintenance du déséquilibre de liaison généralisé sur l'ensemble du génome par la limitation de la recombinaison au niveau de larges régions génomiques (Ollitrault & Navarro, 2012). En effet, l'analyse par la cytométrie en flux de la taille du génome nucléaire des espèces du genre *Citrus* expose une différenciation qui atteint les 10% entre *C. reticulata* (0.74 pg/2C) et *C. medica* (0.81 pg/2C) (Ollitrault *et al.*, 2003) qui représentent les tailles la plus petite et la plus grande du genre *Citrus*. Quant aux espèces secondaires, elles portent un génome de taille intermédiaire reflétant leur origine interspécifique et indiquant le stade avancé atteint par les taxons de base vers la spéciation.

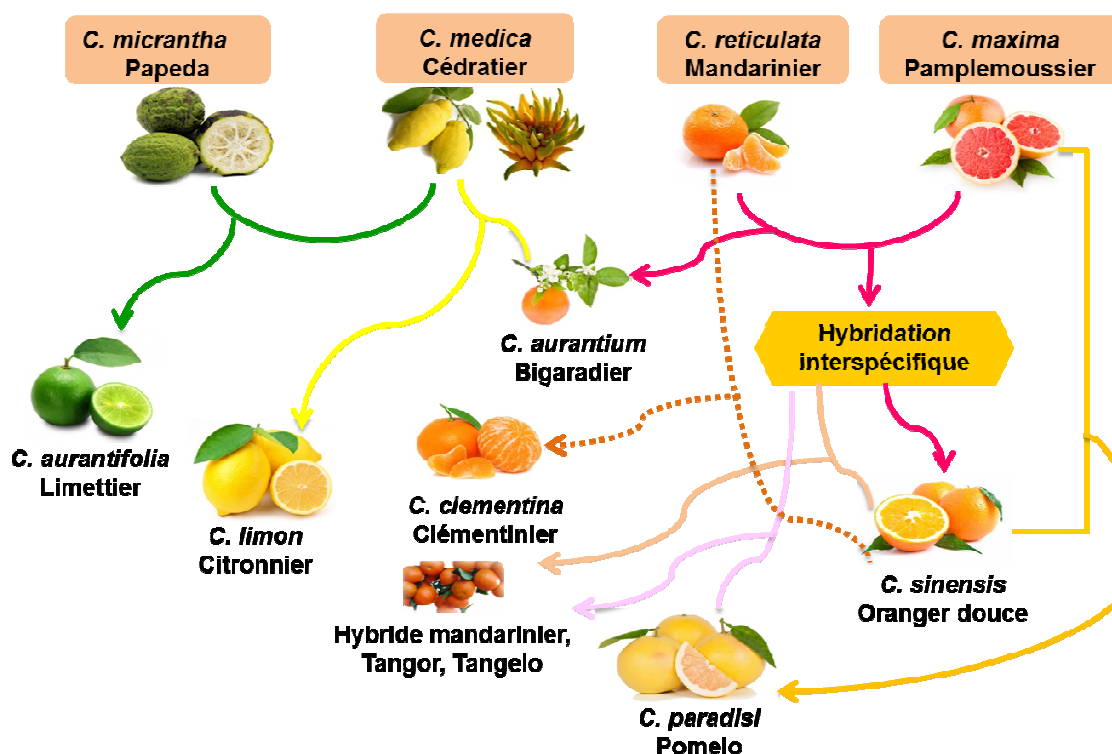


Figure 12 : Représentation schématisque de l'origine de quelques espèces secondaires du genre *Citrus* (Ollitrault *et al.*, 2012 avec modification par Amel Oueslati)

➤ Les bigaradiers

La bigarade (*C. aurantium*, Figure 13) est un orange amère. Les Arabes l'ont introduit en Espagne et en Sicile (Oliva & Valdés, 2003), où elle a été incorporée dans les jardins, les parcs et les rues. Le bigaradier, est un arbrisseau épineux et très décoratif de quatre à cinq m d'hauteur. Il est largement implanté dans la région méditerranéenne. Son tronc est très ramifié et ses feuilles sont de couleur vert brillant. Ses fleurs, très odorantes, sont blanches à l'intérieur et pourpres à l'extérieur. Le fruit du bigaradier est plus petit que celui d'un oranger doux. Il est ovoïde et jaune foncé. Les fruits sont souvent des baies cloisonnées à pulpe vésiculeuse et juteuse formée de poils intracarpellaires : ce sont des hespéridés propres aux agrumes (Hadrich *et al.*, 2008).



Figure 13 : Fruit de l'espèce *Citrus aurantium* (GIF, 2010)

La bigarade est hautement apomictique est utilisée comme porte-greffe pour l'orange douce et d'autre biotypes de *Citrus* (Moore *et al.*, 2001). Son génome est prédominé par la mandarine et introgressé par des gènes du pamplemousse (Barrett & Rhodes, 1976). Selon Nicolosi *et al.* (2000), la bigarade est considérée comme un hybride naturel entre pamplemousse et mandarine. Ceci a été confirmé lors d'une étude récente basée sur une hybridation interspécifique entre les deux parents (Garcia-Lor *et al.*, 2013).

➤ Les oranges douces

L'oranger doux (*C. sinensis*, Figure 14) est un arbrisseau d'environ d'une hauteur de six à quinze mètres avec des branches épineuses. L'écorce est mince, lisse et gris-brun à verdâtre. Les feuilles sont parfumées, de couleur vert foncé, de taille moyenne, pointues au bout, arrondies à la base, avec des pétioles ailés composés à la fois de la pousse et de la tige. Le fruit peut être arrondi ou aplati. Il est d'une baie constitué de cinq à douze loges contenant de nombreuses graines et dérivent d'un ovaire unique (Harley *et al.*, 2006).



Figure 14 : Fruit de l'espèce *Citrus sinensis* 'Maltaise' (GIF, 2010)

L'orange douce dérive d'un parent mâle mandarine et d'une femelle issue d'un back cross entre pamplemousse et mandarine. Toutefois, Garcia Lor *et al.* (2013) et Wu *et al.* (2014) ont montré que de nombreuses régions du génome de l'oranger étaient en homozygotie phylogénomique entre *C. reticulata* et d'autres plus rare *C. maxima*. Les deux parents de l'oranger étaient donc d'origine interspécifique *C. reticulata* et *C. maxima*.

➤ Les citrons

Le citronnier (*C. limon*, Figure 15) est un agrume mesurant de trois à six mètres d'hauteur. Ses feuilles sont ovales, allongées, pointues à l'extrémité et d'un vert pâle virant au rouge à maturité. Il possède des pétioles ailés formant la pousse. Les fleurs sont blanches, violacées à la surface inférieure des pétales. Le fruit est ovale, avec une large papille apicale, composé de six à dix loges à écorce épaisse et pointé de glandes. A maturité, le fruit est de couleur jaune et à pulpe acide (Coile, 1995).



Figure 15 : Fruit de l'espèce *Citrus limon* (GIF, 2010)

Les citrons regroupent les variétés acides et quelques formes douces. Ultérieurement, ils ont été utilisés dans la médecine et maintenant leur utilisation est très répandue dans la gastronomie : jus et arômes (Moore *et al.*, 2001).

Les différentes variétés de citronnier jaunes de type Méditerranéen sont très similaires morphologiquement et présentent des caractères biochimiques communs. Leur diversité est supposée être due à des mutations à partir d'un ancêtre unique d'origine

interspécifique (Gulsen & Roose, 2001; Luro et al., 2008; Uzun & Yesiloglu, 2012). Diverses hypothèses ont été proposées sur l'origine des citronniers et des limettiers. Les citronniers méditerranéens seraient des hybrides naturels entre cédratier et limettier (Scora, 1975 ; Barrett & Rhodes, 1976) ou bien entre cédratier et oranger amère (Nicolosi et al., 2000 ; Gulsen & Roose, 2001). Cette dernière hypothèse a été confirmée par des travaux récents (Ollitrault et al., 2012 ; Garcia Lor et al., 2013 ; Turk et al., 2016).

➤ Les pomelos

Le pomelo (*C. paradisi*, Figure 16) est un agrume à fruits comestibles. Ses feuilles sont persistantes, luisantes, ovales et vertes. L'arbuste mesure entre cinq et six mètres. Il est décoratif et présente une floraison en petites grappes de fleurs blanches parfumées. Souvent confondu avec le pamplemousse, le fruit contient au tant de vitamine C qu'une orange mais deux fois moins sucré.



Figure 16 : Fruit de l'espèce *Citrus paradisi* (GIF, 2010)

Le pomelo est considéré résultat d'une hybridation entre un pamplemoussier et un oranger doux (Barrett et Rhodes, 1976 ; Ollitrault et al., 201; Garcia Lor et al., 2013). La variabilité génétique au sein du groupe des pomelos est très faible. En effet, lors d'une étude, un haut taux de similarité a été observé chez les variétés étudiées (Scora, 1988; Gmitter et al., 1992; Fang et al., 1997; Nicolosi et al., 2000).

➤ Les limes

Les limes appelés aussi les citrons verts (*C. aurantifolia*, Figure 17), incluent des variétés acides et douces avec un niveau de ploïdie différent. En effet, les limes acides représentent : les petits fruits Mexicains (West Indian, Key) qui sont diploïdes ($2n=18$) et les grands fruits de Tahiti (Perse) qui sont triploïdes (Moore et al., 2001). En effet, Barrett & Rhodes (1976) les ont considérés comme des *Citrus* apomictiquement perpétués. Tanaka (1969) a suggéré la Papeda comme parent direct des variétés diploïdes. Les limettiers de type Mexicain sont des hybrides apparentés au cédrat et à la papeda (Scora, 1975 ; Nicolosi et al., 2000) ou bien résulteraient d'hybridations entre trois espèces : cédratiers,

pamplemoussiers et *Microcitrus* (Barrett & Rhodes, 1976 ; Barkley *et al.*, 2006). Les marqueurs de type RFLP, RAPD et SCAR ont considéré les limes diploïdes comme des hybrides directes entre *Citrus micrantha* et *Citrus medica* (Scora 1975 ; Federici *et al.* 1998 ; Nicolosi *et al.* 2000) et les résultats récents (Ollitrault *et al.*, 2012 ; Garcia Lor *et al.*, 2013 et Curk *et al.*, 2016) confirment cette hypothèse. Quant à l'origine parentale des variétés triploïdes n'est connue que pour certaines variétés.



Figure 17 : Fruit de l'espèce *Citrus aurantifolia* (Aujardin, 2016)

➤ Les clémentines

La clémentine (*C. clementina*, Figure 18) est un fruit sans pépin de couleur orangé-rouge pouvant présenter jusqu'à 1/5^{ème} de la surface de l'épiderme en vert. Elle est de calibre moyen, avec un diamètre maximal de la section équatoriale compris entre 46 et 68 mm. La clémentine présente une teneur en jus de 42 % minimum et un goût acidulé (Verzera, 1998). Cette clémentine sans pépin, facile à éplucher, est devenue le type de mandarine le plus apprécié dans les régions méditerranéennes. Elle est surtout cultivée en Espagne, au Maghreb et en Corse (Verzera, 1998). Bien que relativement récente, la clémentine est à l'origine de quelques mutations, chacune d'elles étant caractérisée par l'époque de maturité, la fertilité, la couleur et la consistance du fruit (Barboni, 2006).

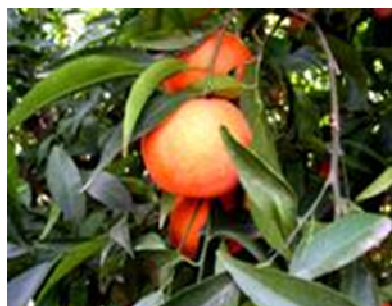


Figure 18 : Fruit de l'espèce *Citrus clementina* 'Marisol' (GIF, 2008)

Elle fut découverte en 1898 par le père Clément à Oran en Algérie (Trabut, 1904). Tanaka (1961), l'érige au rang d'espèce et lui donne le nom de *Citrus clementina*, et Swingle & Reece (1967) la considèrent comme une variété dans le groupe des mandarines

ou comme un hybride plus proche des oranges douces. Certains taxonomistes la considèrent comme un hybride de la mandarine méditerranéenne pollinisée par une variété ornementale d'orange amère. Des travaux de marquage moléculaire tranchent et attestent que la clémentine est le résultat d'un croisement naturel entre un mandarinier (*Citrus reticulata*) et un oranger (*Citrus sinensis*) (Deng *et al.*, 1996; Luro *et al.*, 2001). Des travaux récents ont confirmé cette hypothèse et ont suggéré l'appartenance de la clémentine au groupe horticole des tangors (ollitrault *et al.*, 2012).

➤ Les petits agrumes

Les mandariniers regroupent plusieurs variétés issues d'hybridation ou de mutation. Les variétés issues d'hybridation montrent une variation génétique alors que les variétés mutationnelles montrent une très faible variabilité génétique comme chez le groupe des clémentines (Bretó *et al.*, 2001; Barkley *et al.*, 2006; Uzun & Yesiloglu, 2012). Le groupe des mandarines est constitué par plusieurs variétés hybrides dérivées de mandarine x mandarine (Fortune et kinnow), mandarine x variété secondaire ou mandarine x tangelo.

✓ Les tangors

Les tangors sont des cultivars Australiens capables de produire des fruits sans pépins en absence de pollinisation (Sykes & Possingham, 1992). La tangor est un hybride entre la mandarine et l'orange. Elle regroupe plusieurs variétés comme Ellendale, Murcott, ortanique (Alonso & Howell, 1983), Kiyomi et Temple (Figure 19) (Combrink *et al.*, 2013).



Figure 19 : Fruit du groupe horticole des Tangors 'Temple' Minneola' (C. Jacquemond-INRA, 2012)

✓ Les tangelos

Le fruit de la tangelo (*C. tangelo*, Figure 20) possède neuf à onze loges, il est à écorce mince aromatique, le jus est doux et acidulé (Coile, 1995). La tangelo est un hybride entre la mandarine et le pomelo telles que la tangelo 'Minneola' et 'Orlando' (Swingle & Reece, 1967), elle utilisée aussi comme géniteur pour d'autres hybridations.



Figure 20 : Fruit de l'espèce *Citrus tangelo* 'Tangelo Minneola' (Dave's, 2015)

3. L'information génétique chez les agrumes : Structure, Ploïdie et Mode de transmission

L'information génétique chez les agrumes est stockée au niveau du noyau cellulaire (le génome nucléaire) et au niveau du cytoplasme (le génome chloroplastique et le génome mitochondrial).

Le génome nucléaire, issu d'un brassage de chromosomes maternels et paternels, est à l'origine de la singularité de chaque individu. L'ADN nucléaire présente ainsi plusieurs caractéristiques pouvant compliquer son analyse: l'hétérozygotie, la recombinaison, la duplication de ces loci ...

Le génome chloroplastique contenu dans les chloroplastes, organites cellulaires dérivés des procaryotes. Le chloroplaste résulte d'une endosymbiose entre une cyanobactérie et une cellule eucaryote. Il est très abondant dans le cytoplasme des cellules végétales supérieures vu qu'il est le siège de la photosynthèse. Le chloroplaste porte une molécule d'ADN généralement transmise selon un mode monoparental. L'ADN chloroplastique (ADNcp) est une molécule circulaire bicaténaire, présente à quelques milliers de copies dans chaque cellule végétale et sa taille varie entre 135 et 160 kb chez les angiospermes (Omstead & Palmer, 1994). Cette molécule est généralement subdivisée en trois régions (Palmer, 1985). Une région inversée répétée (IR) et deux régions en copie unique: la petite région simple copie (SSC) et la grande région simple copie (LSC). L'évolution de la taille des génomes chloroplastiques peut être le résultat des processus d'expansion et de contraction des régions IR, et des insertions et délétions dans les séquences non codantes (Gillham, 1994). La région IR varie de 20 à 25 Kb. En effet, la faible taille de l'ADNcp des légumineuses fait suite à la perte d'une copie entière d'une région IR. A l'inverse, les plus grands génomes sont liés à l'augmentation de la taille de la région inversée répétée (Palmer *et al.*, 1987), alors que la SSC et la LSC ont des tailles conservées. Il semble que la présence d'IR dans l'ADNcp stabilise structurellement la

molécule en diminuant la fréquence des réarrangements (Palmer & Thomson, 1982). L'ADNcp possède un taux d'évolution conservateur, il est non recombinaut et présente l'avantage de fournir des polymorphismes exceptionnellement peu homoplasiques pour les études phylogénétiques.

Le génome nucléaire des *Citrus* et des autres genres de la famille des Aurantioideae est diploïde avec un nombre de chromosome $2n=18$ (Krug, 1943). Certains cas d'euploïdie ont été observés chez les *Citrus*. Les variations généralement observées sont des cas de triploïdie et de tétraploïdie. En effet, la 'Kumquat Hong kong' présente un génome tétraploïde tandis que la lime 'Tahiti' présente un génome triploïde (Longley, 1925). Le génome des *Citrus* est de petite taille. La taille de l'haploïde *C. sinensis* est de 372 Mb (Arumuganathan & Earle, 1991). Une étude antérieure sur l'estimation de la taille du génome des *Citrus* en utilisant la cytométrie en flux a montré une variation de taille significative entre les différentes espèces du genre *Citrus* (Ollitrault, 1994).

En effet, le génome le plus grand est celui de *C. medica* alors que le plus petit est celui de *C. reticulata* (avec des valeurs moyennes respectives de 398 Mb et 360 Mb/génome haploïde). Des valeurs intermédiaires ont été enregistrées chez *C. maxima*, *C. aurantium*, *C. paradisi* et *C. Limon* (avec des valeurs moyennes respectives de 383, 368, 381 et 380 Mb/génome haploïde). Quant à l'ADNcp, il est haploïde et transmis de façon uni-clonale (Reboud & Zeyl, 1995). Chez les agrumes, la transmission pourrait être sexuée ou asexuée.

4. Importance économique et distribution géographique

4.1 A l'échelle mondiale

Selon l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture (FAO), la production mondiale des agrumes se situe autour de 135 million de tonnes en 2013. Dont 71.3 MT (58%) d'oranges, 28.66 MT (23%) de mandarines et de petits agrumes (clémentines, tangerines...), 14.94 MT (12%) de limes et de citrons et 8.22 MT (7%) de pamplemousses et de pomelos (FAO, 2013), Figure 21).

Cette production est concentrée dans le continent Asiatique avec une superficie de 60,63 ha. Ainsi, les plus importants pays producteurs d'agrumes sont la Chine (32,57 MT), le Brésil (19,73 MT), les Etats-Unis d'Amérique (10,13 MT), l'Inde (10,09 MT), le Mexique et l'Espagne chacun vers les 7 MT (FAO, 2013).

La production de l'orange, le principal agrume produit dans le monde, est dominée par le Brésil, le premier pays producteur de l'orange dans le monde, avec 17,54 MT par an suivie par les États-Unis d'Amérique, la Chine puis l'Inde avec des valeurs respectives de 7,57 MT, 7,3 MT et 6,42 MT.

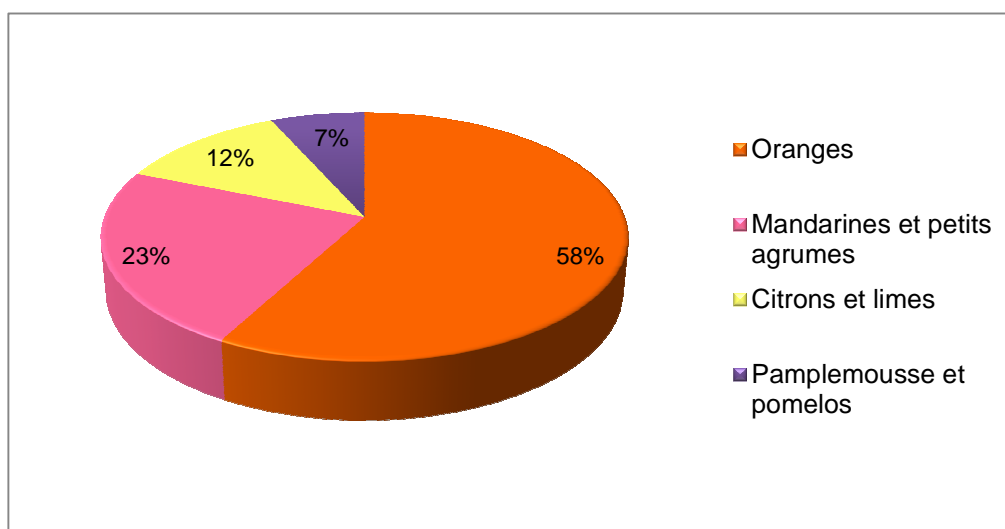


Figure 21 : Répartition mondiale de la production des agrumes par groupe (FAO, 2013)

Selon la FAO (2013), la Chine est la plus grande productrice de la mandarine et des petits agrumes avec 17,17 MT de production annuelle, suivie par l'Espagne avec 2,19 MT. Les limes et les citrons sont produits en première position par l'Inde (2,52 MT) suivie par le Mexique (2,13 MT) la Chine avec 1,94 MT vient en 3^{ème} position suivie en 4^{ème} position par l'Argentine (1,3 MT). La Chine (3,71 MT) et les États-Unis d'Amérique (1,07 MT) sont les deux principaux pays producteurs du pamplemousse et du pomelo dans le monde (FAO, 2013).

Plus que 17 MT sont produites dans la région méditerranéenne par les 10 pays membres (Sections nationales) du Comité de Liaison Des Agrumes Méditerranéenne (CLAM). L'Espagne occupe la première place dans la région méditerranéenne avec une production annuelle d'agrumes totale de 6,37 MT. L'Egypte vient en deuxième position avec 4,09 MT suivie par la Turquie (3,68 MT) (FAO, 2013).

4.2 A l'échelle nationale

L'agrumiculture en Tunisie demeure une des principales richesses de l'arboriculture Tunisienne. Elle vient au troisième rang des productions après la culture des olives et des dattes. La superficie est d'environ de 34,84 ha plantés par plus que 7 millions de pieds de différentes variétés. En effet, la péninsule du Cap Bon, située entre le golfe de Tunis et celui d'Hammamet s'étire vers l'Europe précisément vers la Sicile (à peine 140 Km), représente 71% de cette superficie d'agrumes cultivés et demeure la plus importante région agrumicole. Toutefois, la Tunisie connaît d'autres zones de production avec des superficies de culture d'agrumes différentes: les environs de Tunis (Ben Arous, La Manouba, Mornag, La Soukra...) représentent 10% de la superficie totale. La plaine de Kairouan, la région de Bizerte (Ras Djebel, Elalia...) et la région intermédiaire du Nord (Béja, Bousalem...)

représentent respectivement 6%, 5% et 6% de la superficie totale. Quant aux régions allant de la côte du pays (Bir Bou Regba, Sousse, Mehdia, Monastir) au Sud (Sfax) jusqu'à Djerba, ne représentent que 2% de la superficie cultivée totale avec de petites plantations aux oasis (Figure 22).

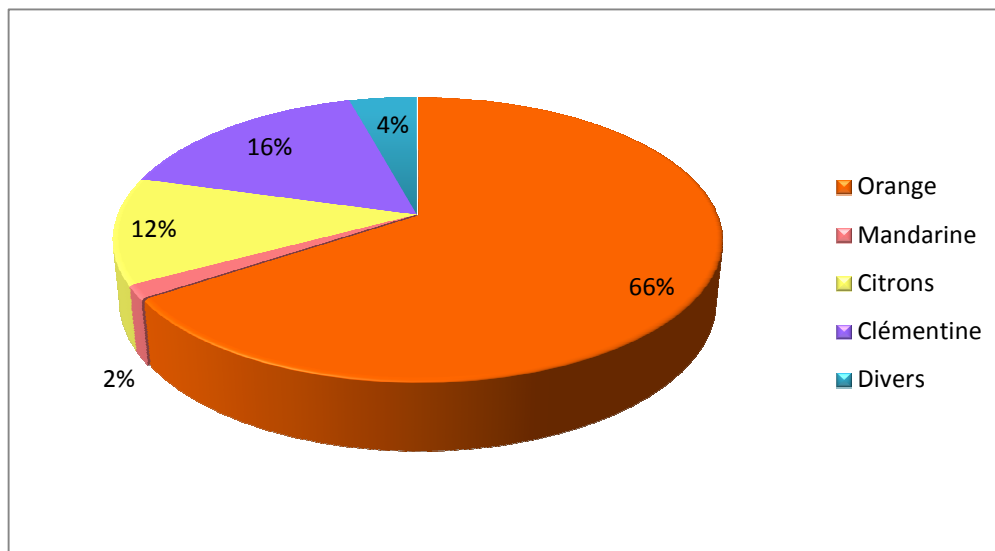


Figure 22 : Répartition des différents groupes d'agrumes selon leur production en Tunisie (GIF, 2016)

Toutes ces régions agrumicoles plus ou moins importantes et réparties sur tout le pays, se situent dans des zones climatiques et édaphiques très différentes, ce qui laisse supposer l'adaptation des différentes variétés cultivées.

Au cours de ces dernières années, la production agrumicole tunisienne vient en 8^{ème} position en Méditerranée avec 431 milles tonnes en moyenne/an pour la campagne de 2014/2015. Une pré-estimation de production pour la campagne de 2015/2016 est de 380 milles tonnes en moyenne/an enregistrant ainsi une possibilité de régression de production de 12% par rapport à l'année précédente (GIF, 2016). En effet, la production de l'orange vient en première position avec une moyenne de 250 milles tonnes par an suivie par les clémentines, les citrons, les diverses variétés et enfin les mandarines avec des valeurs respectives de : 62,5 milles tonnes, 46 milles tonnes, 16 milles tonnes et 5,8 milles tonnes (GIF, 2016) (Figure 23).

L'orange, le principal agrume produit en Tunisie regroupe plusieurs variétés locales et introduites. Les oranges 'Maltaise' sont les plus produites en Tunisie avec une pré-estimation de production de la campagne 2015/2016 de 111,5 milles tonnes suivie par l'orange 'Navel', douce et 'Valencia' avec des valeurs respectives de : 93 milles tonnes, 32 milles tonnes et 16 milles tonnes.

Ainsi, la Tunisie est le seul pays producteur et exportateur mondial de l'orange 'Maltaise demi-sanguine' réputée pour être « la Reine des Oranges ». Cette variété cultivée

sur le territoire de la péninsule du Cap Bon, est unanimement connue et appréciée à l'échelle mondiale (CTA, 2015).

En outre, la production des différents fruits d'agrumes couvre les besoins locaux et dégage un surplus destiné à l'exportation et estimé pour l'année 2016 (données d'exportation jusqu'au 11 Mars 2016) à environ 13 milles tonnes dont 12,26 milles tonnes d'orange 'Maltaise' principalement exportés en France suivie par les citrons avec un taux d'exportation de 405,1 tonnes exportées aux pays du Golf et en Europe. Il est important de mentionner que, la Tunisie exporte 94% de sa production agrumicole en France (GIF, 2016).

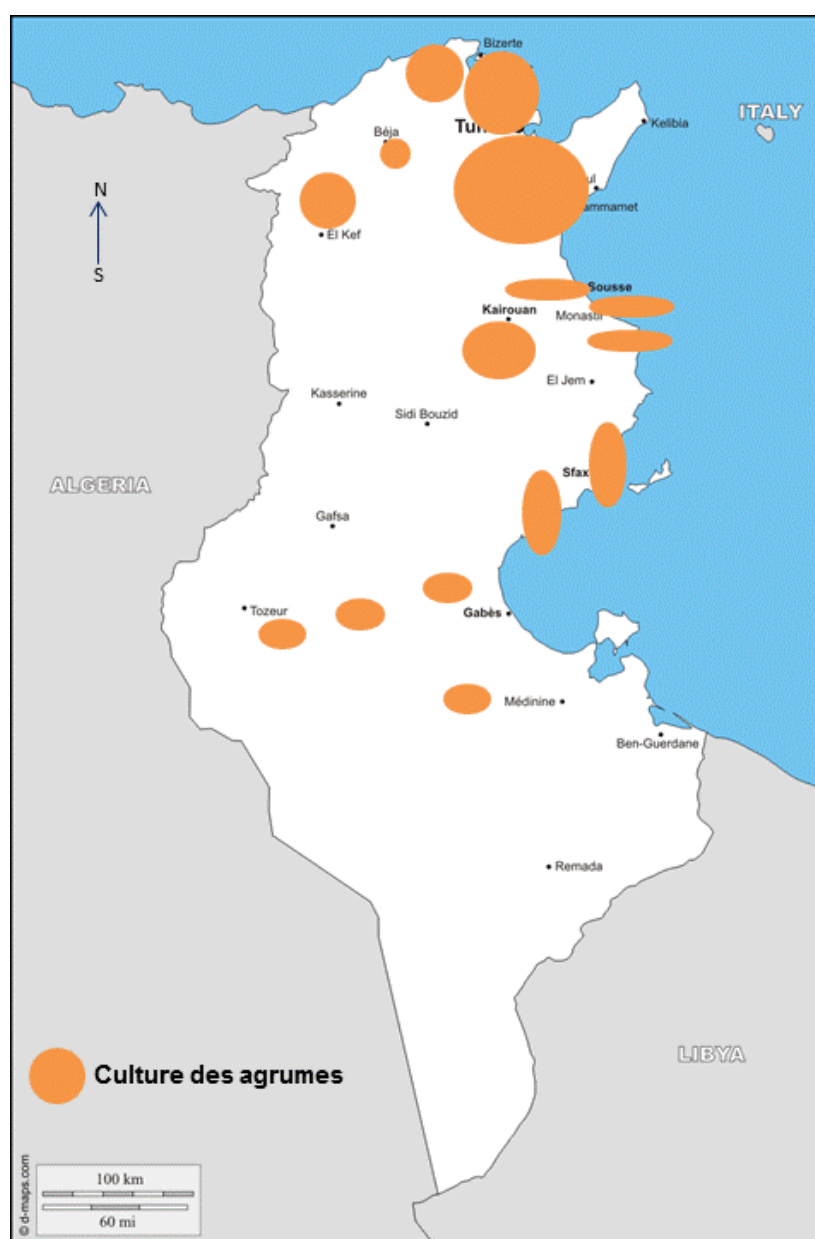


Figure 23 : Carte de l'agrumiculture en Tunisie (Préparée par Amel Oueslati Bahri)

5. Les principales contraintes de l'agrumiculture

La culture des agrumes, première production mondiale, caractérisée par une large diversité du sol et du climat est confrontée à diverses contraintes biotiques et abiotiques selon les régions de production.

En effet, la culture des agrumes est touchée par la salinité (Storeoy & Walker, 1999) et la sécheresse, deux principaux facteurs abiotiques limitant du développement agrumicole. Ainsi, au niveau du bassin Méditerranéen, le climat subtropical favorise la sécheresse ce qui entraîne un déficit hydrique. En Tunisie, la production des agrumes est très liée aux pluviométries. Entre autre, la sécheresse entraîne la chute des feuilles (Tudela & Primo-Millo, 1992), la diminution du potentiel hydrique et de la conductance stomatique (Gomez-Cadenas *et al.*, 1996) ainsi que la baisse du rendement et de la qualité des fruits (Yakushiji *et al.*, 1998 ; Jacquemond *et al.*, 2013). D'autre part, au niveau du bassin Méditerranéen, les agrumes cultivés sur des sols calcaires et basiques entraînent la Chlorose ferrique qui touche 20 à 50 % des arbres (Jaeger *et al.*, 2000). Les sols calcaires qui dominent le centre de la Tunisie engendrent une déficience en fer des agrumes et les sols trop acides peuvent, de même, toucher l'agrumiculture (Ollitrault & Navarro, 2012 ; Jacquemond *et al.*, 2013). La Tunisie, surtout dans les zones agrumicoles du Cap Bon, du Centre et du Sud, est caractérisée, entre autre, par un sol salin dû à des causes naturelles géologiques et/ou induit par l'activité de l'homme liée à l'irrigation ce qui affecte la croissance et la productivité des plantes cultivées. Récemment, avec le réchauffement climatique qui induit l'élévation accélérée du niveau de la mer, les terres agricoles et la qualité de l'eau sont ainsi affectées.

D'autres parts, l'agrumiculture est affectée par les nématodes, les champignons, les bactéries, les phytoplasmes, les spyroplasmes, les virus et les viroïdes (Ollitrault & Navarro, 2012). Ces stressés biotiques affectent la qualité du fruit et causent une réduction de la production. Certaines maladies sont présentes dans la majorité des régions comme celles causées par : les nématodes, l'oomycète *Phytophthora* sp. ou le virus de la Tristeza (CTV : *Citrus Tristeza Virus*) ce qui empêche d'une part l'utilisation de certains porte-greffes unanimement connus excellent et d'autre part, réduisent la production et la qualité des variétés spécifiques (Ollitrault & Navarro, 2012). De même, la Tunisie est confrontée à la maladie de brindilles sèches (Malsecco) et à la Tristeza. Ainsi, certains ravageurs se sont limités à une zone de culture spécifique, et d'autres se sont propagés rapidement dans de nouvelles zones. Par ailleurs en Tunisie, le puceron *Toxoptera citricida* considéré comme le vecteur le plus efficace du virus qui, en effet, contribue à aggraver sa propagation. Il est actuellement au Portugal, en Espagne et récemment au Maroc (GIF, 2016). D'autre part, en Tunisie, le viroïde *Citrus exocortis* a touché les arbres des clémentiniers (Elleuche *et al.*, 2003) et le viroïde *Hop stunt viroid* (HSVd) a touché les arbres *Citrus* (Hamdi *et al.*, 2011). La

maladie du chancre des agrumes est causée par la bactérie *Xanthomonas axonopodis* pv. Citri (Hasse) (Ollitrault & Navarro, 2012). Cette maladie a détérioré plusieurs cultures d'agrumes en Amérique du Sud et en Asie, mais récemment, a reconnu une large propagation incontrôlée à la Floride des Etats Unies d'Amérique (USA). La chlorose variégée causée par la bactérie *Xylella fastidiosa* (Ollitrault & Navarro, 2012) ainsi que la maladie de la mort subite causée par *Citrus sudden death associated virus* ont endommagé la production agrumicole au Brésil. En Afrique, une maladie fongique causée par *Phaeomalaria angolensis* (De Carvalho & O. Mendes P.M. Kirk) appelée la Cercosporiose, détériore les arbres des citrons et constitue une contrainte pour les porte-greffes (Ollitrault & Navarro, 2012).

6. L'amélioration des agrumes

6.1 Principaux objectifs

Les chercheurs visent à développer le patrimoine fruitier agrumicole selon une stratégie d'amélioration bénéfique à ce secteur. Ainsi, différentes approches ont été mise en place : l'amélioration des porte-greffes et l'amélioration variétale par voie sexuée ou récemment génétique.

En Tunisie, l'analyse du germoplasme local des porte-greffes d'agrumes était une étape préliminaire avant de mettre en œuvre les programmes d'amélioration des agrumes dans le but d'identifier les porte-greffes adaptables à chaque région du pays. Les porte-greffes jouent un rôle clef dans la résistance à certaines maladies, l'adaptation des variétés aux conditions pédoclimatiques et leurs performances agronomiques (en qualité et en quantité). Pour ce, un contrôle de la conformité génétique des collections des porte-greffes de Sbikha et d'El Gobba et une analyse de la diversité génétique des porte-greffes régionaux ont été réalisés sur 206 accessions de porte-greffes d'agrumes (Snoussi *et al.*, 2006). Le bigaradier est le porte-greffe le plus utilisé en Tunisie (plus de 90%) et dans toutes les régions agrumicoles mondiales. Il lutte contre la *Phytophthora* mais il présente en association avec les différentes variétés une sensibilité à certaines maladies graves qui menacent le secteur telles que la maladie de brindilles sèches (Malsecco) et la Tristeza. Pour ce faire, une mise en place de programmes de sélection de nouveaux porte-greffes dans le monde entier a été menée. En Tunisie, de nouveaux porte-greffes alternatifs du bigaradier ont été proposés aux producteurs et aux pépiniéristes : *Citrance carrizo*, *Citrumelo* Swingle, *Citrus Volkameriana* et *Citrance C-35* (CTA, 2016). En prenant compte du climat des régions agrumicoles tropicales et tempérées et de ses persécutions sur la culture des agrumes, des porte-greffes nanisant, qui limite la vigueur des arbres, ont été sélectionnés. Un autre facteur d'amélioration vise à produire des porte-greffes à graines

polyembryonnées permettant sa propagation par apomixie partielle (Ollitrault & Navarro, 2012).

D'autre part, l'amélioration variétale est un objectif majeur de l'agrumiculture. Elle permet de s'adapter au marché selon les conditions environnementales de la région de production en visant à améliorer les qualités pomologiques et organoleptiques du fruit. L'amélioration variétale des agrumes s'est intéressée à d'autres critères de sélection tel que: l'absence de pépin, la stérilité mâle et femelle, (l'oranger Navel) et l'auto-incompatibilité (le clémentinier). Actuellement, de nouvelles stratégies sont adoptées dans les projets de création variétale. Ainsi, la Tunisie a suivi un assainissement des variétés anciennes et une introduction de nouvelles variétés afin de viser des dates de maturité échelonnées ce qui permet d'enrichir les étalages des variétés tardives et donc étaler la période de production offrant, de même, une production variée et de bonne qualité.

L'amélioration de la qualité nutritionnelle des fruits fait également partie des nouveaux objectifs à l'échelle internationale (Ollitrault & Navarro, 2012). D'autre part, l'amélioration variétale vise à augmenter la résistance aux maladies tels que la maladie du dragon jaune (*Huanglongbing*) et le chancre citrique (Cuenca *et al.*, 2013).

6.2 Atouts et facteurs limitants

Les stratégies d'amélioration conventionnelle des agrumes, basées sur la recombinaison sexuée, sont très limitées. En effet, la majorité des espèces secondaires des agrumes sont apomictiques et les embryons d'origine nucellaire, ainsi formés, limitent ou empêchent le développement de l'embryon zygotique. Ceci limite le redressement d'une large population sexuée, ce qui a poussé les chercheurs à éviter d'utiliser les variétés apomictiques comme parent femelle dans plusieurs programmes d'amélioration. Plusieurs variétés (comme l'orange 'Navel' et la mandarine 'Satsuma') à pollen et ou à ovule stérile ne sont pas, de même, utilisées dans les programmes d'amélioration. L'auto-incompatibilité (comme chez les clémentiniers, les tangors, les tangelos et les pamplemoussiers), l'inter-incompatibilité de certaines variétés auto-incompatibles et l'inter-incompatibilité unilatérale sont fréquentes chez plusieurs espèces, également, limitent la possibilité de sélectionner des parents pour des croisements spécifiques (Ollitrault & De Rocca Serra, 1992). Chez la plupart des espèces d'agrumes, la phase juvénile est longue. Au moins 5 ans sont obligatoires pour initier la floraison et se débarrasser des épines, caractère indésirable associé à la juvénilité. Ces facteurs biologiques et la difficulté de gérer le domaine d'évaluation à larges progénitures et à schémas d'amélioration multi générationnel, sont les principales raisons pour le succès relativement faible des programmes d'amélioration conventionnelle (Ollitrault & Navarro, 2012).

La stratégie d'amélioration des agrumes, se basant sur la propagation végétative, a eu recours à l'induction du polymorphisme dans un seul cycle permettant ainsi de développer des sélections clonales. Cette sélection est faite sur des mutations spontanées identifiées chez des génotypes obtenus par des hybridations, la mutagenèse induite ou après avoir eu recours aux approches biotechnologiques comme l'hybridation somatique et la transformation génétique. Ces différentes méthodes ont été développées afin de résoudre le problème constaté lors des programmes d'amélioration conventionnelle des agrumes. Dans le but de prouver l'efficacité de ces stratégies, le transfert des importants traits commerciaux des parents à la descendance doit être optimisé. Ceci dépend d'une gestion systématique de l'hétérozygotie et à travers la connaissance de la ségrégation du caractère sélectionné. La mise en œuvre d'une sélection assistée par les marqueurs est particulièrement importante pour tous les traits qui n'ont pas été phénotypés durant la phase juvénile (Ollitrault & Navarro, 2012).

Parmi les taxons de base, les pamplemoussiers et les mandariniers ont un important polymorphisme inter-cultivars aux niveaux moléculaire et morphologique. L'amélioration intra-spécifique peut être effectuée par la recombinaison sexuée. Cependant, quand l'objectif est de diversifier une gamme variétale autour d'un idéotype précis (par exemple la clémentine), le recours à l'hybridation sexuée est empêché par une structure génomique parentale fortement hétérozygote. Des méthodes permettant la modification ponctuelle du génome comme la mutagenèse ou la variation somaclonale, l'hybridation somatique asymétrique ou la transformation génétique doivent être appliquées. Tenant compte de leurs structures génétiques, les mêmes approches seront utilisables dans l'amélioration des espèces comme chez l'orange douce et le pomelo, si leurs caractères spécifiques méritent d'être maintenus (Ollitrault & Navarro, 2012).

Suite à l'évolution des techniques biotechnologiques utilisées dans les programmes d'amélioration génétique des agrumes, telles que : la mutagenèse (dirigée et induite), la sélection de mutants spontanés, la recombinaison sexuée, la polyploïdie, les cybrides et la transgénèse, une reconstruction des structures génomiques hybrides à partir d'une combinaison des génotypes des taxons ancestraux pourrait être réalisable.

7. Application des marqueurs génétiques chez les Agrumes

L'avancée de la biologie moléculaire au cours de ces dernières décennies a permis d'innover les techniques d'isolement d'ADN, la généralisation de la PCR et du séquençage et de fournir des connaissances sur les génomes des plantes. Les travaux concernant la phylogénie et la diversité génétique se sont basés sur l'étude des génomes cytoplasmiques (chloroplastique et mitochondrial) et nucléaires. Ainsi, les potentialités d'un génome ou d'une

région particulière dépendent des facteurs extérieurs et intrinsèques tels que : la taille, l'abondance, l'accessibilité, le niveau taxonomique considéré, le mode d'évolution et le niveau d'homoplasie.

7.1 Evolutions des techniques moléculaires et étude de la diversité génétique

L'étude de la diversité génétique se base sur l'information morphologique, biochimique et moléculaire. Les études biochimiques sur les *Citrus* remontent aux années 70 avec l'utilisation des isoenzymes (Torres *et al.*, 1978) pour identifier les différentes variétés d'agrumes. Dans les années 90, ces études ont été approfondies suite à la découverte des marqueurs du polymorphisme de l'ADN. Les principales études de la diversité génétique des *Citrus* basées sur les marqueurs moléculaires ont ciblé l'amplification aléatoire de l'ADN polymorphe [RAPD : *Random Amplified Polymorphie DNA* ; (Luro *et al.*, 1994 ; Biswas *et al.*, 2010; Pessina *et al.*, 2011; Malik *et al.*, 2013; Tripolitsiotis *et al.*, 2013; Jung *et al.*, 2005)], le polymorphisme d'amplification inter-microsatellites [ISSR : *Inter-Simple Sequence Repeat* ; (Fang *et al.*, 1997 ; Gulsen & Roose, 2001a; Biswas *et al.*, 2010; Yang *et al.*, 2010; Uzun *et al.*, 2011a; Lombardoa *et al.*, 2012; Kumar *et al.*, 2013; Santos *et al.*, 2013; Tripolitsiotis *et al.*, 2013; Şahin-Çevik & Moore, 2012; Kim *et al.* 2013)], le polymorphisme de longueur des fragments de restriction [RFLP : *Restriction Fragment Length Polymorphisms* ; (Federici *et al.*, 1998 ; Ramadugu *et al.*, 2013)], le polymorphisme de longueur des fragments amplifiés [AFLP : *Amplified Fragment Length Polymorphism* ; (Liang *et al.*, 2007 ; Al-Sadi *et al.*, 2012; Zhou *et al.*, 2013; Bernardi *et al.*, 2010; Fang *et al.*, 2010; Kepiro & Roose, 2010; Li *et al.*, 2010; Distefano *et al.*, 2011; Licciardello *et al.*, 2011; Xiao *et al.*, 2011; Kacar *et al.*, 2012)], le polymorphismes d'amplification inter-rétrotransposons [IRAP : *Inter-Retrotransposon Amplified Polymorphism* ; (Fang et Roose, 1997; Breto *et al.*, 2001; Sanker & Moore, 2001; Ahmad *et al.*, 2003; Barkley *et al.*, 2006; Pang *et al.*, 2007 ; Biswas *et al.*, 2010; Santos *et al.*, 2013) ainsi que l'amplification de séquences reliées à des marqueurs polymorphes [SRAP : *Sequence-Related Amplified Polymorphism* ; (Li & Quiros, 2001)], et les séquences polymorphes amplifiées et coupées [CAPS : *Cleaved Amplified Polymorphic Sequences* ; (Omura *et al.*, 2000; Jiang Dong *et al.*, 2010; Amar *et al.*, 2011; Yamamoto *et al.*, 2011; Shimada *et al.*, 2014)].

Les études des marqueurs moléculaires déjà cités ont montré un excès en hétérozygotie pour certaines espèces ancestrales, en particulier les mandarines et les pamplemousses. Contrairement aux citrons et aux oranges trifoliés qui présentent des niveaux variables de diversité génétique. Quant aux retrotransposons, ils jouent un rôle majeur dans la diversité des agrumes (Breto *et al.*, 2001).

Plusieurs groupes de chercheurs ont produit des marqueurs de séquences exprimées (EST : *Expressed Sequence Tag*) chez des différentes espèces d'agrumes (Forment *et al.*,

2005; Tero *et al.*, 2006). Plusieurs banques de chromosomes artificiels bactériens (BAC) d'agrumes et de leur proches parents ont été développées (Yang *et al.*, 2001). Des marqueurs des séquences simples répétées (SSR : *Simple Sequence Repeat*) ont été développés à partir de bibliothèques génomiques (Kijas *et al.*, 1997; Barkley *et al.*, 2006; Froelicher *et al.*, 2008): 56 SSR ont été obtenus à partir de banque de données EST (Chen *et al.*, 2006) et 200 SSR ont été développés (Luro *et al.*, 2008) de 1600 séquences microsatellites de 37 000 EST identifiés par Terol *et al.* (2007). De même, plus que 7600 SSR ont été identifiés à partir de séquences BAC *end* (Terol *et al.*, 2008). De même des InDels ont été développés à partir des séquences BAC *end* (Ollitrault *et al.*, 2012).

Les marqueurs moléculaires de type SSR ont été appliqués dans des études de la diversité génétique des germoplasmes d'agrumes (Nematollahi *et al.*, 2009; Uzun *et al.*, 2012b 2011b dans référence biblio; Kijas *et al.*, 1995; Barkley *et al.*, 2006; Uzun *et al.*, 2011b; Garcia-Lor *et al.*, 2013; Novelli *et al.*, 2000; Roose *et al.*, 2000; Corazza-Nunes *et al.*, 2002; Liu *et al.*, 2002; Ahmad *et al.*, 2003; Golein *et al.*, 2005; Jiang Dong *et al.*, 2006; Novelli *et al.*, 2006; Caruso *et al.*, 2008; Chen *et al.*, 2008; Luro *et al.*, 2008; Ghorabaie *et al.*, 2010; Gulsen *et al.*, 2010; Ollitrault *et al.*, 2010; Amar *et al.*, 2011; Biswas *et al.*, 2011; Cristofani-Yaly *et al.*, 2011; El-Mouei *et al.*, 2011; Kamiri *et al.*, 2011; Ollitrault *et al.*, 2011; Singh *et al.*, 2011; Biswas *et al.*, 2012; Garcia-Lor *et al.*, 2012; Golein *et al.*, 2012; Ollitrault *et al.*, 2012a; Polat *et al.*, 2012; Snoussi *et al.*, 2012; Chai *et al.*, 2013; Kacar *et al.*, 2013; Liu *et al.*, 2013; Yildiz *et al.*, 2013; Hou *et al.*, 2014).

De même des InDels, ont été utilisés dans la caractérisation du germoplasme d'agrumes et dans l'analyse de la diversité génétique intra et interspécifique (Ollitrault *et al.*, 2012 ; Snoussi *et al.*, 2012).

La technologie moléculaire n'a pas cessé d'évoluer. Ainsi, les BAC *end* étaient aussi utilisés dans le développement de marqueurs de polymorphisme de nucléotide simple (SNP : *Single Nucleotide Polymorphisms*). Terol *et al.* (2008) ont identifié 6617 SNPs putatifs à partir de séquences BAC *end* de *C. clementina*, dont 622 ont été transmis à l'ensemble du genre par la méthode *GoldenGate array* (Ollitrault *et al.*, 2010). Deng *et al.* (2010) ont définis des SNPs à partir de l'orange douce et à partir des données EST de la mandarine 'Satsuma'. Il est important de noter que le développement des SNPs (fait à partir de l'orange douce) remonte à l'année 2004 par Novelli *et al.*

Cependant, les SNPs sont des variations naturelles qui ne concernent qu'un seul nucléotide dans une séquence donnée comparée à un ou plusieurs individus. Ils peuvent être individuellement responsables d'un caractère spécifique ou peuvent représenter une variation neutre qui pourrait être utile dans l'évaluation de la diversité dans un contexte évolutif (Collins, 1998 ; Brookes, 1999). Il s'agit d'un marqueur moléculaire codominant et sa

détection nécessite une connaissance préalable de la séquence à analyser (Brookes, 1999). Chez les végétaux, les SNPs ont rapidement remplacé les SSRs en amélioration et en génétique végétale. Le recours aux marqueurs SNPs est très répandu car ils sont abondants, répartis d'une façon homogène dans tout le génome, efficaces, plus stables et potentiellement facile à identifier que les microsatellites.

Les marqueurs SNPs peuvent, aussi, être automatisés avec un rapport de coût/efficacité raisonnable (Duran *et al.*, 2009). Ainsi, une fois le SNP est identifié, un génotypage par la PCR compétitive allèle spécifique (KASPar) pourrait être effectué. Une autre technique utilisée pour le génotypage est le polymorphismes de conformation simple brin (SSCP : *Single Strand Conformation Polymorphism*) (Olivares-Fuster *et al.*, 2007; Simsek *et al.*, 2011) et ce, en se basant sur les différences dans la conformation de la structure tertiaire de l'ADN. Cette technique permet d'analyser un grand nombre d'échantillon mais nécessite des conditions d'essai rigoureuses.

Ainsi, avec l'évolution des techniques de séquençage, il est devenu très utile chez les *Citrus* d'utiliser les SNPs dans : l'analyse de la variation génétique et génomique, comme le mappage de l'ensemble du génome, marquage des traits importants, comparaison de l'évolution génomique, l'identification et la classification des cultivars de divers taxons, l'étude de l'association génotype/phénotype et la sélection assistée par marqueur (Botstein & Risch, 2003; Morales *et al.*, 2004; Xing *et al.*, 2005; Lijavetzky *et al.*, 2007; Ollitrault *et al.*, 2012b).

Cependant, des études de diversité au niveau du genre *Citrus* et des genres proches ont été menées en utilisant les marqueurs SNPs (Novelli *et al.*, 2004; Jiang Dong *et al.*, 2010; Ollitrault *et al.*, 2012b; Bernardi *et al.*, 2013; Chen & Gmitter, 2013; Cuenca *et al.*, 2013a; Distefano *et al.*, 2013; Fujii *et al.*, 2013; Garcia-Lor *et al.*, 2013; Olivares-Fuster *et al.*, 2007; Luro *et al.*, 2011; Simsek *et al.*, 2011; Cuenca *et al.*, 2013b; Ye JunLi *et al.*, 2010; Kaçar *et al.*, 2014).

7.2 Applications à l'étude des taux d'apomixie

La polyembryonie est un processus de développement d'embryons dérivés d'un tissu mère où la graine donnent plusieurs plantules et ce, lors de la germination. Les embryons ainsi développés peuvent avoir une origine zygotique : issus d'un gamète mâle et d'un gamète femelle. Ou bien une origine nucellaire. Il s'agit d'un type d'apomixie produisant des plants asexuellement : sans fécondations et sans méiose.

Dans les cultivars polyembryonnés des agrumes, les embryons somatiques développés à partir d'une cellule de nucelle (le tissu nourricier de l'ovule) du plant maternel ont un profil génétique identique au pied mère (Kobayashi *et al.*, 1967 ; 1979). Le nombre des embryons produits par graine pour des cultivars d'agrumes importants économiquement

varie de 2 (citron et lime) à 20 (pomelo, orange douce, satsuma, kumquat, rough lemon et orange trifolié) et à plus que 20 chez la mandarine 'Ponkan'. La polyembryonie limite strictement l'utilisation de ces cultivars dans les croisements, mais il s'agit d'un trait utile dans la stabilité des portes-greffes et dans leur propagation à partir de la graine, assurant ainsi des plants sains uniformes dans le verger.

Des études antérieures basées sur des tests de croisement suggèrent qu'un ou plusieurs gènes dominants sont impliqués dans la détermination de la polyembryonie chez les agrumes (Hong *et al.*, 2001 ; Iwamassa *et al.*, 1967 ; Parlevliet & Cameron, 1959). Plusieurs loci de la polyembryonie ont été révélés chez *Citrus volkameriana* et *Poncirus trifoliata* (deux portes-greffes) (Asins *et al.*, 2002). Plusieurs marqueurs moléculaires voisins des loci majeurs de la polyembryonie chez la mandarine 'Satsuma' ont été identifiés (Nakano *et al.*, 2006). Cinq marqueurs AFLPs associés au locus de la polyembryonie chez *Poncirus trifoliata* ont été identifiés (Kepiro & Roose, 2009). Nakano *et al.* (2008) ont identifié des CAPS et des SNPs arrondissant le locus de la polyembryonie. A partir desquels, ils ont évalué la structure haplotypique et construit des contigs BAC des haplotypes spécifiques couvrant la région génomique du locus de la polyembryonie. En 2012, Nakano *et al.* ont identifié plusieurs ORF aidant à identifier le ou les gènes associés à la polyembryonie chez les *Citrus*. En 2013, Nakano *et al.* ont identifié un gène *msg-2* suggérant qu'il pourrait être associé à la polyembryonie chez les *Citrus*.

D'autre part, la polyembryonie zygotique a été révélée chez les génotypes des *Citrus* (Aleza *et al.*, 2010). Les marqueurs RAPDs ont été utilisés pour le repérage de plantules zygotiques dans un semis de mandarines (Bastianel *et al.*, 1998), de bigarades (Rao *et al.*, 2007; 2008) et de 'Rough lemon' (Savita *et al.*, 2012). De même, des SNPs ont permis la discrimination de zygotes dans des semis (Zhu *et al.*, 2013).

Actuellement, les chercheurs sont toujours intéressés par l'étude de la polyembryonie, trait caractéristique des *Citrus*. En effet, des analyses ont été effectuées en utilisant des marqueurs RAPDs pour étudier les deux origines possibles de la polyembryonie (Mondal & Saha, 2014) et les marqueurs SSRs pour étudier son origine nucellaire (Mondal *et al.*, 2015).

7.3 Application aux études phylogénétiques

La reconstruction phylogénétique facilite la compréhension des processus évolutifs qui génèrent la diversité biologique. Au cours de ces dernières décennies de nombreux marqueurs ont été utilisés pour l'étude des relations phylogénétiques: des marqueurs morphologiques, enzymatiques, biochimiques et récemment moléculaires (Darlu & Tassy, 1993).

Le choix du type de marqueur moléculaire est crucial dans l'établissement des relations phylogénétiques entre les groupes à étudier. Ces marqueurs doivent posséder un

taux d'évolution compatible avec le niveau de différenciation des espèces étudiées et doit de plus fournir des données aisément analysables.

7.3.1 Phylogénie chloroplastique

Chez les végétaux, l'ADN chloroplastique est le plus utilisé dans l'étude des relations phylogénétiques. En effet, la petite taille de l'ADNcp, son abondance dans la cellule végétale, son hérédité monoparentale et surtout son taux d'évolution conservateur, font du génome chloroplastique un outil privilégié pour les études phylogénétiques (Omstead & Palmer, 1994 ; Cleg *et al.*, 1994). Cependant, l'ADNcp peut contourner les problèmes des relations phylogénétiques compliquées au sein des *Citrus* et des genres proches qui sont liés aux hybridations, l'apomixie et la polyploïdie (Nicolosi *et al.*, 2000). En fonction du niveau phylogénétique et de la différenciation des espèces étudiées, plusieurs types d'approches moléculaires sont envisageables pour mettre en évidence les variations de l'ADNcp. Les chercheurs ont eu recours aux données de séquence des régions non-codantes de l'ADNcp pour évaluer la phylogénie évolutive de la sous famille des Aurantioideae. Ces régions non-codantes ont, en effet, tendance à évoluer plus rapidement que les régions codantes, par accumulation d'insertions et de délétions.

Cependant, des études antérieures sur la sous famille des Aurantioideae utilisant des données de séquences chloroplastiques de l'intergène *atpB-rbcL* et de l'intron *rps16*, ont révélé un polymorphisme plus ou moins informatif pour révéler de solides déduction phylogénétiques (Samuel *et al.*, 2001). Quant aux études phylogénétiques sur la tribu des Citreae basées sur des séquences des régions *trnL-trnF* et *rps16* n'ont pas été efficaces dans la résolution des liens de parenté au niveau des tribus et des sous-tribus appartenant à la sous famille des Aurantioideae (Samuel *et al.*, 2001; Araujo *et al.*, 2003 ; Morton *et al.*, 2003).

Les études phylogénétiques se sont approfondies davantage et plusieurs chercheurs ont essayé de clarifier la phylogénie au sein du groupe des agrumes vrais, parmi lesquelles l'équipe de Bayer *et al.* (2009). Ainsi, neuf gènes ont été utilisés et ont révélé une phylogénie proche à celle proposée par Swingle & Reece, (1967). Une autre étude sur le groupe des agrumes vrais basée sur trois régions chloroplastiques *trnL-trnF*, *psbH-petB* et *trnS-trnG*; (Lu *et al.*, 2011) confirme la monophylie du groupe et montre sa subdivision en six genres confirmant ainsi les résultats de Swingle & Reece (1967). Récemment, Carbonnel Caballero *et al.* (2015) ont révélé une structure phylogénétique du groupe des agrumes vrais en analysant le polymorphisme de séquences chloroplastiques issues des données de re-séquençage du génome complet positionné sur le génome chloroplastique de référence de l'orange douce (Bausher *et al.*, 2006). Ceci a révélé trois clades essentiels: le premier regroupe les cédrats et les espèces Australiennes, le second associe les pamplemousses

avec *C. micrantha* et le troisième associe les mandarines avec la papaya *C. ichangensis*. Quant aux *Poncirus* et *Fortunella*, ils apparaissent séparément.

D'autre part, l'analyse des régions non-codantes de l'ADNcp permet également d'étudier les phylogénies à des niveaux taxonomiques encore plus bas. En effet, Les espaceurs intergéniques de l'ADNcp sont des régions potentiellement utilisées dans les études de la systématique et de la phylogénie chez les différents taxons végétaux y compris les *Citrus* (Araujo *et al.*, 2003). Yingzhi *et al.* (2007) ont utilisé l'intergène *trnL-trnF* pour clarifier la phylogénie des mandarines sauvages et cultivées issues de la Chine. Cependant, cet intergène était inefficace dans la détermination des liens de parenté à des hauts niveaux taxonomiques et montre son efficacité dans les études phylogénétiques à des bas niveaux taxonomiques. De même, Jung *et al.* (2005), se sont basés sur cette région pour élucider les liens de parenté entre les *Citrus* Coréens. Les chercheurs ont eu recours à d'autres intergènes comme la région *rbcL-ORF 106* et la région *trnD-trnT*. Une étude sur les *Citrus* de l'Inde, a été effectuée sur les intergènes précédemment décrits. Cette étude a montré l'importance de l'intergène *trnL-trnF* dans la détection des variations génétiques chez les génotypes de *Citrus* indiens. D'autres études se sont basées sur des séquences de gènes, des espaceurs et des introns chloroplastiques comme le gène *matK* (Penjor *et al.*, 2013), l'espaceur *trnL-trnF* (Li *et al.*, 2007; Jena *et al.*, 2009); les régions *rbcL-ORF106*, *trnL-trnF* et *trnF-trnVr* (Yamamoto *et al.*, 2013), *psbH-petB*, *trnL-trnF*, le gène *rbcL* (Penjor *et al.*, 2010) et la région *trnS-trnG* (Li *et al.*, 2010).

En plus des données des séquences, plusieurs études ont analysé la phylogénie maternelle du genre *Citrus* en utilisant différentes approches moléculaires. Ainsi, la plus ancienne étude était basée sur les marqueurs chloroplastiques de type RFLP (Olmstead & Palmer 1994; Green *et al.*, 1986 Yamamoto *et al.*, 1993). Des SSRs chloroplastiques ont été développés (Deng *et al.*, 2006; Garcia *et al.*, 2015; Li *et al.*, 2015; Curk *et al.*, 2016) ainsi que les marqueurs moléculaires CAPSs (Lotfy *et al.*, 2002; Yamamoto *et al.*, 2013), ils étaient efficaces et très utiles dans l'étude de la phylogénie maternelle du genre *Citrus*. Toutes ces études ont permis de différencier entre les quatre taxons de base du genre *Citrus* (*C. reticulata*, *C. maxima*, *C. medica* et *C. micrantha*) et éclaircir leur contribution dans la phylogénie maternelle des espèces secondaires.

7.3.2 Phylogénie nucléaire

Plusieurs types de marqueurs nucléaires ont été développés et utilisés dans des études phylogénétiques. Des marqueurs dominants comme les RAPDs (Naz *et al.*, 2014), les ISSRs (Fang *et al.*, 1997 ;1998; Gulsen & Roose, 2001a ; 2001b) et les AFLPs (Pang *et al.*, 2007; Biswas *et al.*, 2011; Nartvaranant, 2011; Pessina *et al.*, 2011; Zhang *et al.*, 2011) ont été largement utilisés dans les études phylogénétiques basées sur le génome nucléaire

des *Citrus* et pour lesquels, les informations sur les séquences génomiques n'ont pas été disponibles. Des marqueurs codominants plus puissants comme les SSRs (Gulsen & Roose, 2001a; Luro *et al.* 2001) ont été aussi développés et utilisés chez les *Citrus*. Des InDels et des SNPs ont été utilisés pour éclaircir les liens de parenté entre les *Citrus* (Garcia-Lor *et al.*, 2013).

7.4 Application à la cartographie génétique

Depuis les années 80, la cartographie de liaison a été utilisée mais sans avoir recourt à des cartes de haute densité avec des marqueurs de séquences définies (Durhem *et al.*, 1992; Jarrell *et al.*, 1992). Une carte génétique du génome de deux agrumes a été élaborée à l'aide des marqueurs moléculaires RFLPs et RAPDs et des isoenzymes (Luro *et al.*, 1995). En effet, la taille de la population et le nombre mappé de marqueurs codominants demeurent généralement faibles. Un effort collaboratif fourni pour développer des marqueurs codominants (spécialement des SSRs) et établir une carte, a néanmoins été récompensé par la construction de cartes pour *Citrus sinensis* et *Poncirus trifoliata* (Chen *et al.*, 2007). De même, des séquences ont été utilisées afin de développer des marqueurs moléculaires de type SSRs pour une cartographie génétique comparée des espèces de *Citrus* (Ollitrault *et al.*, 2010).

Le recours à la mise en évidence de marqueurs moléculaires de type SNP suggérerait leur utilité dans la cartographie génétique comparative chez les *Citrus* et dans l'étude d'association génétique chez *C. reticulata*. Ainsi, leur abondance dans le génome permet de créer des cartes génétiques à très haute résolution (Ollitrault *et al.*, 2012a). En effet, les promesses de l'étude étaient réalisées et une carte génétique de la clémentine a été développée permettant ainsi de faciliter l'assemblage chromosomique de la première séquence de référence du génome des *Citrus* (Ollitrault *et al.*, 2012b).

De même des InDels ont été développés à partir des séquences BAC *end* dans le but d'être utilisés dans la cartographie génique, la caractérisation du germoplasme d'agrumes et l'analyse de la diversité génétique intra et interspécifique (Ollitrault *et al.*, 2012). Par la suite, en associant les données ESTs et SSRs cartographiés et en se basant sur la carte génétique de la clémentine une carte génétique du pamplemousse a été développée (Chai *et al.*, 2013).

Récemment, une cartographie génétique des centromères des neuf chromosomes de la clémentine a été effectuée. En effet, la cartographie de la localisation des centromères permet de fournir des informations essentielles dans l'analyse de la structure génétique puisque la ploïdie de la plante est affectée par la distance génétique du centromère (Aleza *et al.*, 2015).

7.5 Application à l'étude du déterminisme génétique des caractères (QTL)

Vue l'hétérozygotie élevée chez les agrumes, l'étiquetage des gènes a été largement utilisée. Celle-ci a été utilisée pour identifier les différences entre les individus au niveau de plusieurs traits particuliers tels que le nanisme (Cheng & Roose, 1995), la tolérance au virus de tristeza des agrumes (Fang *et al.*, 1998 ; Gmitter *et al.*, 1997 ; Mestre *et al.*, 1997), la tolérance des agrumes aux nématodes (Ling *et al.*, 2000) et les embryons nucellaires (Kepiro, 2004).

Pour repérer les gènes affectant les différents traits, une analyse de locus de caractère quantitative (QTL : *Quantitative Trait Loci*) a été appropriée. Des régions contenant des gènes qui influencent la tolérance à la salinité (Tozlu *et al.*, 1999), et la résistance à la Tristeza des agrumes ont été identifiées (Asins *et al.*, 2004 ; 2011). Des QTL liés à la polyembryonie ont été révélés (Asins *et al.*, 2002) ainsi que des QTL liés à la qualité des fruits et de la résistance à des maladies (Takita, 2015).

8. Progrès récents dans le domaine de la génomique chez les Agrumes

8.1 Les séquences de référence complète du génome

Parmi les objectifs majeurs du Consortium International de la génomique des agrumes (ICGC : *International Citrus Genome Consortium*) est de produire des séquences de haute qualité des différents génomes des agrumes. Le séquençage des génomes des agrumes permet de faciliter la comparaison des génomes des herbacées et des ligneuses pérennes et de produire des ressources valables pour étudier des questions biologiques d'une importance critique permettant ainsi l'amélioration génétique des agrumes (Talon & Gmitter, 2008).

Pour minimiser les complications associées à une forte hétérozygotie du génome, le consortium International ICGC a caractérisé trois haploïdes. Le choix final était sur la clémentine 'Nules' (*C. clementina*) pour son statut exempt d'agents (Aleza *et al.*, 2009). Pour ses qualités, la séquence du génome complet haploïde de la clémentine sert comme une première référence génomique pour toutes les espèces du genre *Citrus* et ses genres proches. Dès 2014, cette séquence est disponible sur le site *Phytozome.com* et sur la base de données du consortium ICGC.

En parallèle du projet haploïde, un second projet du génome diploïde de l'orange douce en utilisant la technologie du séquençage à haut débit de deuxième génération, appelée le pyroséquençage (454 Roche), a été élaboré. Le CIGC a sélectionné un diploïde d'orange douce 'Ridge Pineapple' comme un génome cible pour le projet de séquençage. L'orange douce a été sélectionnée pour sa place prépondérante dans l'agrumiculture du monde entier et le 'Ridge Pineapple' a été sélectionné spécifiquement pour sa production de

fruit à pépins caractérisés par la présence de minimum de réarrangements chromosomiques qui sont principalement associés aux fruits sans pépins (Gmitter *et al.*, 1992). La comparaison des deux séquences (clémentine et orange douce) a montré que la séquence de l'orange douce est beaucoup plus fragmentée que celle de la clémentine. Toutefois, une similarité importante dans le nombre de détections des gènes prédits au niveau de chaque génome montre une bonne moyenne de l'espace de couverture des gènes codants dans les deux assemblages (Gmitter *et al.*, 2012).

Cependant, une collaboration entre une équipe chinoise et une équipe américaine a permis de réaliser un séquençage de bonne qualité de l'orange douce (*C. sinensis*), principal génome de référence (Xu *et al.*, 2013). Il s'agit d'un séquençage du génome complet, vu cet important exploit, un premier jet diffus à la communauté scientifique a précédé la copie finale (Figure 24 ; Xu *et al.*, 2012).

Bien que le séquençage des génomes de différents groupes d'agrumes a été élaboré par Shimizu *et al.* (2012) et ce en séquençant les mandarines 'Satsuma' et 'Ponkan' ainsi que différents hybrides, le séquençage du génome complet diploïde de l'orange douce et celui de l'haploïde de la clémentine restent les principales références utilisées par la communauté scientifique des agrumes.

D'autre part, il convient de mentionner que le recours au séquençage du génome complet nucléaire n'a pas été la seule orientation des chercheurs. En effet, la séquence du génome complet chloroplastique de l'orange (*C. sinensis*) et celle de la lime 'Omani' (*C. aurantiifolia*) ont été publiées respectivement par Bausher *et al.* (2006) et Su *et al.* (2014).

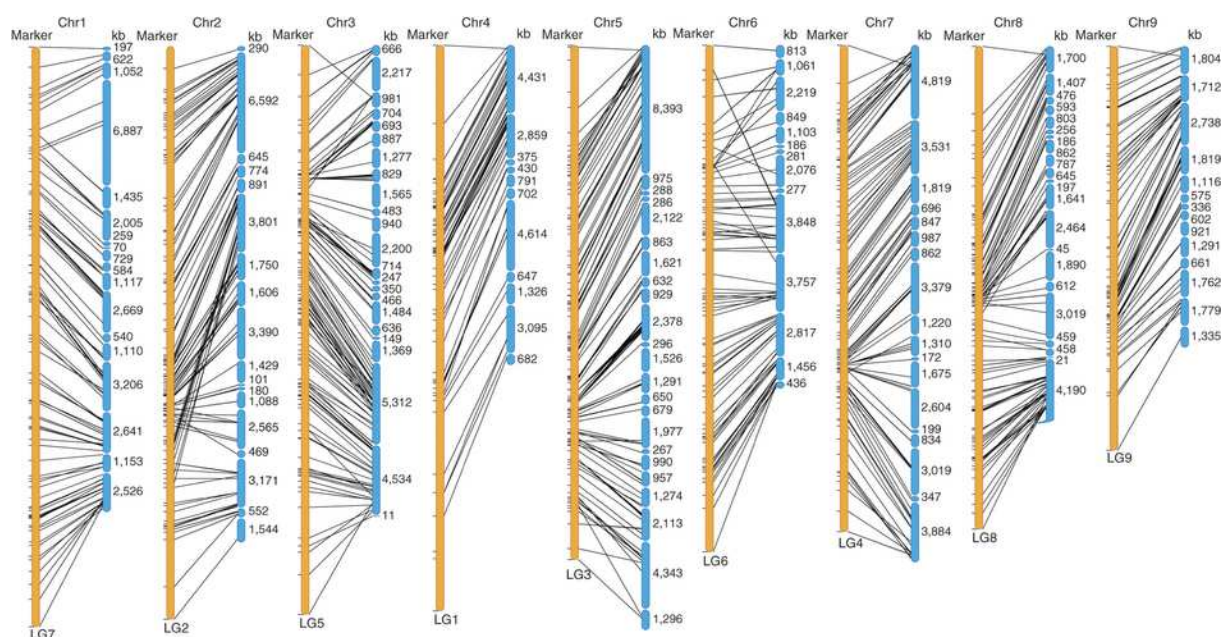


Figure 24 : Assemblage des scafolds avec les marqueurs génétiques correspondants chez l'espèce *C. sinensis* (Xu et al., 2012)

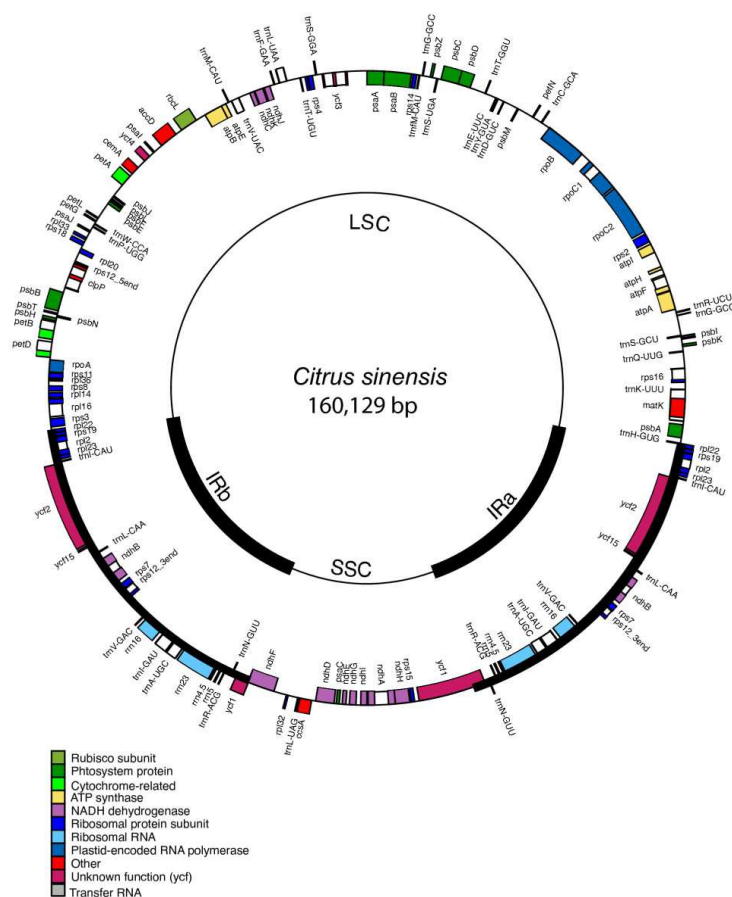


Figure 25 : Carte génétique circulaire du génome chloroplastique de l'espèce *C. sinensis* (Bausher et al., 2006)

Gènes à l'extérieur du cercle : transcrits dans le sens des aiguilles d'une montre

Gènes à l'intérieur du cercle : transcrits dans le sens inverse des aiguilles d'une montre

En effet, le génome chloroplastique de l'orange douce séquencé par Bausher *et al.* (2006) est le génome référence utilisé par la communauté scientifique (Figure 25). Ainsi, il présente une taille de 160,129 pb. Le génome contient une paire de la région IR de 26,996 pb séparée par les régions SSC et LSC avec des tailles de 18,393 pb et 87,744 pb respectivement. Ce génome contient 133 gènes dont 113 sont en simple copie alors que 20 sont dupliqués dans la région inversée répétée. 81 gènes codent pour des protéines dont 9 sont dupliqués dans la région IR. Il contient 4 gènes codant pour des ARNr, 30 ARNt dont 7 sont dupliqués dans la région IR. 70 gènes sont composés par des introns, parmi lesquels 14 contiennent un seul intron, tandis que 3 (*clpP*, *rps12*, *ycf3*) sont formés par deux introns. Le génome comporte 92.59 % de protéines, 1.74% d'ARNt et 5.65% de gènes codants pour des ARNr. La teneur en GC et en AT chez le génome chloroplastique de *C. sinensis* est de l'ordre de 38.48% et 61.52% respectivement. La teneur en AT est similaire à celles du Tabac (62.2%), du riz (61.1%) et du Maïs (61.5%). La teneur en AT au niveau des régions LSC et SSC est de 63.19 et 66.66%, respectivement, alors qu'au niveau de la région IR elle est égale à 57.05% due à la présence de groupe de gènes codant pour des ARNr (Bausher *et al.*, 2006).

8.2 Re-séquençage et décryptage des structures génomiques interspécifiques

Le consortium international s'est focalisé sur le séquençage des agrumes. Ainsi, une multitude de données de re-séquençage complet de différentes espèces d'agrumes ont été publiées et sont disponibles sur la base des données du centre national de l'information biotechnologique (NCBI : *National Center of Biotechnology Information*).

Les séquences des génomes complets sont devenues une référence pour l'étude de la génomique chez les agrumes. En effet, sachant que le séquençage génomique et l'analyse de la diversité ont permis de mieux comprendre l'histoire de domestication des différentes cultures fruitières comme le cas des agrumes, le séquençage et l'analyse génomique de différentes collections de pamplemoussiers, de mandariniers et d'orangers ont été élaborés pour tester l'hypothèse pamplemousse/mandarine et révéler l'origine de plusieurs importants cultivars d'agrumes. Cette étude vise à décrypter les structures génomiques interspécifiques et ce, en comparant les espèces déjà citées à la référence haploïde de la clémentine. En effet, deux variétés de mandarines ('Ponkan' et 'Willowleaf') généralement considérées comme de vraies *C. reticulata*, ont montré une introgression par *C. maxima*. De même, les génomes de la clémentine, l'orange douce et la bigarade présentent des introgressions par *C. maxima* (Wu *et al.*, 2014). Ces résultats ont été confirmés par la suite par une étude génomique de Curk *et al.* (2015). Ils ont révélé des introgressions de *C. maxima* dans la majorité des mandarines primitives et des mandarines récemment sélectionnées. Ainsi, ils suggèrent que les processus de reticulation et d'introgression *C. reticulata* x *C. maxima* sont

importants dans la domestication des mandarines comestibles. En effet, la large constitution phylogénomique entre *C. reticulata* et *C. maxima* révélée chez les mandarines, les oranges douces, les bigarades, les pomelos et orangelos sont favorables aux études d'association basées sur la structure phylogénomique du germoplasme (Curk *et al.*, 2015).

D'autre part, le consortium espagnol (CITRUSEQ / CITRUSGENN) possède des séquences non publiées de plus de 150 variétés d'agrumes dont l'ancrage a également été réalisé à partir de l'assemblage de l'haploïde de clémentinier de l'ICGC (Terol *et al.*, 2012).

8.3 Puce à ADN et étude de l'expression du génome

L'importance de la technologie des Puces à ADN dans la génomique fonctionnelle, augmente exponentiellement dans pratiquement tout le monde végétal, et particulièrement dans les cultures agricoles. Chez les agrumes, les premiers profils de transcription ont été présentés par Shimada *et al.* (2005) qui a construit des puces à ADNc pour suivre l'expression de l'ARNm de 2213 gènes durant le développement floral. Depuis lors, plusieurs plateformes de puces à ADN spécifiques des agrumes ont été développées. Le consortium génomique espagnol a développé la première génération de puces à ADNc contenant 12672 sondes correspondantes à 6875 unigènes putatifs de la collection de 22 000 ESTs (Formet *et al.*, 2005). De denses puces à ADN (spécifique des agrumes) ont été développées par Martinez-Gody *et al.* (2008) et composées de 24 000 éléments de puces à ADNc contenant 2000 unigènes. Aux Etats Unies d'Amérique, Close *et al.* (2006) ont désigné 30 264 séries de sonde pour l'analyse de l'expression des gènes.

L'application des puces à ADN sur des profils d'expression géniques a été faite par Fujii *et al.* (2007, 2008) et Terol *et al.* (2007). En 2006, Affymetrix a développé et réalisé des 'Genechip' spécifiques des agrumes dont les deux tiers ont été désignés pour l'analyse de l'expression génique utilisant 30264 séries de sonde. Puis, un génotypage de 3219 gènes en utilisant 5023 SNPs identifiés à partir de données de ESTs de *C. sinensis* et d'autres espèces du genre *Citrus* a été élaboré sur le reste des 'Genechip'. Ces chips ont permis la détection de plusieurs pathogènes et surtout ont permis la détection, chez le *Poncirus trifoliata*, de la région génomique contenant l'allèle de la résistance à la CTV. Une autre étude sur la résistance au chancre chez la kumquat (*Fortunella spp*) a été développée en utilisant une puce avec 2254 éléments (Khalaf *et al.*, 2007).

8.4 Les promesses de la GBS pour un décryptage des structures génomiques étendues à l'ensemble des collections et des populations recombinantes

L'avènement du séquençage de la nouvelle génération (NGS : *Next Generation Sequencing*) a suscité une révolution dans la biologie et ce, par l'obtention d'énormes données de séquences brutes. Une nouvelle technique de génotypage a été développée

appelée « Génotypage par séquençage » (GBS : *Genotyping By Sequencing*). Elle est désormais appliquée sur des espèces très diversifiées ainsi que sur des génomes de grande taille (Elshire *et al.*, 2011).

Par ailleurs, une combinaison de la puissance de la technique de séquençage de nouvelle génération avec une représentation réduite du génome qui focalise les ressources de séquençage sur la fin des fragments de restriction, a permis de génotyper rapidement un nombre record d'échantillons. En effet, cette approche utilise une stratégie de réduction de la complexité du génome pour réduire le coût et simplifier l'identification des marqueurs SNPs. En outre, la réduction du coût de la GBS est due aussi à son fonctionnement en multiplexe. En plus, la méthodologie GBS est caractérisée par ses composants clés : sa capacité d'identifier et de génotyper des SNPs simultanément dans les populations d'intérêt et avec un simple protocole, sa production d'un génotypage à haute densité chez la majorité des espèces végétales, sa simplicité relative, sa robustesse, moins d'étapes de PCR et de purification, la digestion enzymatique est indépendante de la taille, se fait avec ou sans séquence de référence et des codes barre efficaces (Davey *et al.*, 2011). Ce qui a rendu cette approche utilisable chez plusieurs espèces végétales comme le maïs, le blé, l'orge, le riz, la pomme de terre et le manioc (Van Poecke *et al.*, 2013).

En outre, la GBS a le potentiel d'être une vraie technologie révolutionnaire dans l'arène de la génomique du monde végétal. Ainsi, cette technique a été appliquée dans différents domaines de recherche allant de l'identification des gènes jusqu'à la génomique assistée utilisée dans les programmes d'amélioration (Thomson *et al.*, 2012 ; Beissinger *et al.*, 2013). L'utilisation de la GBS inclut des applications dans l'identification des marqueurs moléculaires, la phylogénie, les QTL, la cartographie, l'amélioration des plantes cultivées et création variétale (He *et al.*, 2014).

D'autre part, en se basant sur les données du séquençage du génome complet des agrumes, un important nombre de marqueurs SNPs diagnostiques a été révélé permettant la différenciation de *C. maxima* et *C. reticulata*. Une analyse de leurs structures en mosaïque ainsi que celle des espèces secondaires a été effectuée. Cette analyse des structures génomiques interspécifiques a permis de clarifier la phylogénie et faciliter l'amélioration génétique (Wu *et al.*, 2014). Récemment, une étude basée sur la technique du pyroséquençage a déterminé la structure génomique interspécifique des espèces secondaires du genre *Citrus* (Curk *et al.*, 2015). En effet, différentes approches ce sont révélées efficaces, ont été utilisées pour étudier les structures en mosaïque des *Citrus*. Ainsi, l'approche GBS a été utilisée pour re-séquencer des collections de lignées recombinantes afin d'analyser et cartographier différents caractères d'intérêt dans les programmes de sélection spécifiques (Deschamps *et al.*, 2012). En effet, l'approche GBS a montré

d'énormes résultats sur le plan génomique, elle serait par ailleurs prometteuse dans le décryptage des structures génomiques en mosaïque des *Citrus*.

Matériel & Méthodes

Extraction de l'ADN cellulaire total

Le matériel végétal consiste en de jeunes feuilles prélevées sur des agrumes adultes. Ces feuilles ont été lyophilisées pour leur conservation et leur utilisation pour l'extraction des acides nucléiques en utilisant le protocole QIAcube (Qiagen, 2003) et le Mini kit « DNeasy Plant » (Qiagen S.A.; Madride, Espagne).

La qualité et l'intégrité de l'ADN génomique obtenu ont été testées par électrophorèse analytique sur gel d'agarose à 1% (Sambbrook *et al.*, 1989). Ce gel contient une solution de bromure d'éthidium.

La concentration des ADNs obtenus a été estimée à l'aide du Nanodrope ND-1000 version 1. La technique utilise l'absorbance de l'acide nucléique à une longueur d'onde de 260nm. Ainsi, un ADN pur bicaténaire possède un rapport de densité R compris entre 1,8 et 2 avec :

$$R = \frac{DO_{260}}{DO_{280}}$$

Une valeur de R inférieure à 1,8 détecte une contamination protéique. Alors qu'une valeur supérieure à 2 signifie une contamination par les sels, dans ce cas un lavage par l'éthanol est nécessaire.

1. Matériel végétal et séquences utilisées

1.1 Analyse de l'espaceur intergénique *trnL*_(UAA)-*trnF*_(GAA)

27 variétés appartenant à 7 espèces différentes du genre *Citrus* ont été ciblées pour l'analyse de la diversité génétique. L'échantillonnage comprend deux bigarades, quatre citrons, trois pomelos, trois oranges sanguines, cinq oranges Navel, une orange blonde, quatre clémentines, quatre mandarines et une tangelo. L'ensemble du matériel végétal est originaire de la collection du Centre Technique des Agrumes (CTA) et du Groupement Interprofessionnel des Fruits (GIF) de la station de recherche de Zaouiet Jdidi région du Cap Bon de la Tunisie (Tableau 1).

Tableau 1 : Liste des variétés du genre *Citrus* classées selon la classification de Swingle & Reece (1967)

Espèce	Groupe horticole	Cultivar	Origine
<i>C. aurantium</i> L.	Bigarade : Orange amère	Bigaradier du Maroc	GIF
		Bigaradier	GIF
<i>C. limon</i> (L.) Burm.F.	Citron	Feminello	CTA
		Monachello	CTA
		Santa Téréza	CTA
		Interdonato	CTA
<i>C. paradisi</i> Macf.	Pomelo	Pomelo Ruby	GIF
		Pomelo Marsh	GIF
		Pomelo Star Ruby	CTA
<i>C. sinensis</i> (L.) Osbeck	Orange sanguine	Tarocco	CTA
		Maltaise demi sanguine	CTA
		Maltaise balerin	CTA
	Orange blonde	Valencia Late	CTA
	Orange Navel	Lane late	CTA
		New Hall	CTA
		Navel Late	CTA
		Navelina	CTA
		Washington navel	CTA
<i>C. reticulata</i> Blanco (Swing.) (<i>C. clementina</i> hort. Ex Tan.)	Clémentine	Cassar	CTA
		Marisol	CTA
		Hernandina	CTA
		MA3	CTA
<i>C. reticulata</i> Blanco	Mandarine	Ortanique	CTA
		Fortune	CTA
		Encore	CTA
		Nova	CTA
<i>C. x tangelo</i> J. Ingram & H. E. Moore	Hybride Tangelo	Minneola	CTA
<i>C. maxima</i> (Burm.) Merr.			NCBI
<i>C. medica</i> L.			NCBI
<i>Atalantia Monophylla</i> DC.	Fruit proche des agrumes		NCBI

Nous avons adopté la classification botanique de Swingle & Reece (1967) pour les noms latins. Le nom latin de *C. clementina* est adopté par la classification de Tanaka (1977).

Pour des études ultérieures de séquence, nous avons jugé utile d'associer sous forme de séquence relevées de la banque de données NCBI deux espèces du genre *Citrus*, *C. maxima* (pamplemousse) et *C. medica* (cédrat), ainsi qu'un outgroup (*Atalantia monophylla* (L.)) (Tableau 1).

1.2 Analyse basée sur les marqueurs moléculaires SNPs associée au génotypage par la PCR compétitive allèle-spécifique (KASPar)

1.2.1 Sélection des séquences chloroplastiques

Huit régions chloroplastiques utilisées par Bayer *et al.* (2009) dans l'étude de la phylogénie de la sous famille des Aurantioideae ont été extraites de la banque de données NCBI (Annexe 1). Il s'agit de la région codante *atpB*, l'espaceur *rbcl-atpB*, l'espaceur *rps16*, la région *trnL-F*, l'espaceur *rps4-trnT*, l'espaceur *matK-5'trnK*, l'espaceur *psbM-trnD^{GUC}* et l'intron *trnG*. 79 séquences des accessions de la sous famille des Aurantioideae ont été utilisées et l'espèce *Ruta graveolens* L. de la famille des Rutaceae a été choisie comme « outgroup ». Tous les genres de la sous famille des Aurantioideae ont été représentés sauf le genre *Limnocitrus*. En effet, la tribu des Clauseneae a été représentée par trois sous tribus : Clauseniae (sept accessions), Merrilliinae (une accession) et Micromelinae (une accession). La tribu des Citreae a été représentée également par trois sous tribus : triphasilineae (dix accessions), Balsamocitrinae (sept accessions) et Citrinae (cinquante et une accessions). Il est important de noter que nous avons sélectionné toutes les espèces connues de tous les genres de la sous tribu des Citrinae (Tableaux 2 et 3).

1.2.2 Matériel végétal utilisé pour l'analyse des données KASPar

Cent huit variétés appartenant à la sous famille des Aurantioideae comprenant quatre sous tribus : Balsamocitrinea, Citrineae, Triphasilineae (de la tribu des Citreae) et Clauseniae (de la tribu des Clauseneae) ont été choisies pour l'application des marqueurs SNPs sélectionnés. Dans cette étude, nous avons adopté la classification de Swingle & Reece (1967). La sous tribu des Clauseniae comprend trois genres qui sont *Clausena* (trois accessions), *Glycosmis* (une accession) et *Murraya* (deux accessions). La sous tribu des Merrilliinae est représentée par un seul genre *Merrillia* avec une seule accession. La sous tribu des Balsamocitrinea est représentée par sept genres: *Swingle* (une accession), *Aegle* (une accession), *Afraegle* (deux accessions), *Aeglopsis* (une accession), *Balsamocitrus* (une accession), *Feronia* (une accession), et *Feroniella* (une accession). La sous tribu des Triphasilineae inclue cinq genres: *Oxanthera*, *Pamburus*, *Paramignya*, *Triphasia*, et *Wenzelia* avec une accession représentée par chaque genre.

Tableau 2 : Liste des espèces et des accessions de la sous famille des Aurantioideae utilisées dans l'étude *in silico* et KASPar expérimentale

Tribu	Sous tribu		Genre	In silico		KASPar	
				NE	NA	NE	NA
Clauseneae	Micromelinae		<i>Micromelum</i>	1	1	0	0
	Clauseniae		<i>Clausena</i>	2	2	3	3
			<i>Glycosmis</i>	3	3	1	1
			<i>Murraya</i>	2	2	2	2
	Merrilliinae		<i>Merrillia</i>	1	1	1	1
Citreae	Triphasiinae		<i>Luvunga</i>	1	1	0	0
			<i>Merope</i>	1	1	0	0
			<i>Monanthocitrus</i>	1	1	0	0
			<i>Oxanthera</i>	2	2	1	1
			<i>Pamburus</i>	1	1	1	1
			<i>Paramignya</i>	2	2	1	1
			<i>Triphasia</i>	1	1	1	1
			<i>Wenzelia</i>	1	1	1	1
	Balsamocitrinae		<i>Aegle</i>	1	1	1	1
			<i>Aeglopsis</i>	1	1	1	1
			<i>Afraegle</i>	1	1	1	2
			<i>Balsamocitrus</i>	1	1	1	1
			<i>Feronia</i>	1	1	1	1
			<i>Feroniella</i>	1	1	1	1
			<i>Swinglea</i>	1	1	1	1
	Citrinae	Les fruits proches des agrumes	<i>Atalantia</i> ,	3	3	4	4
			<i>Citropsis</i>	2	2	4	4
		Les fruits des agrumes primitifs	<i>Hesperethusa</i>	1	1	1	1
			<i>Pleiospermium</i>	1	1	1	1
			<i>Severinia</i>	1	1	2	2
			<i>Burkillanthus</i>	1	1	0	0
		Les fruits des agrumes vrais	<i>Citrus</i>	17	31	12	58
			<i>Clymenia</i>	1	1	1	1
			<i>Eremocitrus</i>	1	1	1	1
			<i>Fortunella</i>	3	3	6	6
			<i>Microcitrus</i>	6	6	8	8
			<i>Poncirus</i>	1	1	1	2

NE : Nombre des espèces ; NA : Nombre des accessions

Tableau 3 : Liste des espèces et des accessions du groupe des agrumes vrais utilisées dans l'étude *in silico* et KASPar expérimentale

Genre	Espèce	<i>In silico</i>		KASPar	
		NE	NA	NE	NA
<i>Citrus</i>	<i>C. maxima</i>	1	1	1	8
	<i>C. medica</i>	1	1	1	6
	<i>C. ichangensis</i>	1	2	1	2
	<i>C. reticulata</i>	1	2	1	24
	<i>C. aurantifolia</i>	1	3	1	3
	<i>C. aurantium</i>	1	4	1	2
	<i>C. limon</i>	1	3	1	6
	<i>C. paradisis</i>	1	2	1	2
	<i>C. sinensis</i>	1	1	1	2
	Autres	8	12	3	3
<i>Clymenia</i>	<i>Clymenia polyandra</i>	1	1	1	1
<i>Eremocitrus</i>	<i>E. glauca</i>	1	1	1	1
<i>Fortunella</i>	<i>F. hindsii</i>	0	0	1	1
	<i>F. crassifolia</i>	0	0	1	1
	<i>F. japonica</i>	1	1	1	1
	<i>F. margarita</i>	1	1	1	1
	<i>F. obavata</i>	0	0	1	1
	<i>F. polyandra</i>	1	1	1	1
<i>Microcitrus</i>	<i>M. australis</i>	1	1	1	1
	<i>M. australisica</i>	1	1	1	1
	<i>M. inodora</i>	1	1	1	1
	<i>M. garrowayae</i>	1	1	1	1
	<i>M. papuana</i>	1	1	1	1
	<i>M. virgata</i>	0	0	1	1
	<i>M. warburgiana</i>	1	1	1	1
	<i>M. australisica</i> x <i>C. mitis</i>	0	0	1	1
<i>Poncirus</i>	<i>P. trifoliata</i>	1	1	1	2

NE : Nombre des espèces ; NA : Nombre des accessions

Au sein de la sous tribu des Citrineae trois groupes ont été représentés : les fruits primitifs des agrumes, les fruits proches des agrumes et les agrumes vrais. Les fruits primitifs des agrumes incluent les genres *Severinia* (deux accessions), *Pleiospermum* (une accession) et *Hesperthusa* (une accession), les fruits proches des agrumes comprennent deux genres *Citropsis* (quatre accessions) et *Atalantia* (quatre accessions) et le groupe des agrumes vrais représente six genres : *Fortunella* (six accessions), *Eremocitrus* (une accession), *Poncirus* (deux accessions), *Clymenia* (une accession), *Microcitrus* (huit accessions) et *Citrus* (cinquante-huit accessions) (Tableau 2 et 3). Le matériel végétal est issu de la collection gérée par l'Institut National De la Recherche Agronomique (INRA) et le Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) San Giuliano, France, l'institut de recherche agricole de Valence (IVIA : *Instituto Valenciano de Investigaciones Agrarias*) et du service de recherche agronomique (ARS : *Agricultural Research Service*) qui fait partie du département de l'agriculture des Etats Unies d'Amérique (USDA : *United States Department of Agriculture*) (Annexe 2).

1.3 Analyse basée sur la Technique de Génotypage par Séquençage (GBS)

55 variétés ont été sélectionnées de la collection du Centre de Ressources Biologiques (CRB) Citrus géré par l'INRA et le CIRAD (San Giuliano ; France ; Tableau 4). L'échantillonnage comprend des variétés des groupes horticoles mandarinier (onze variétés) et pamplemoussier (six variétés), sélectionnés pour identifier les marqueurs diagnostiques de la différenciation *C. maxima*/*C. reticulata*. Trente-huit variétés supposées dérivées d'une combinaison de ces taxons ont été introduites dans cette étude : cinq mandariniers hybrides issus de programmes d'amélioration Nord-Américain (*C. reticulata*), dix-huit tangelos (*C. reticulata* x *C. paradisi*) et hybrides tangelos, deux orangelos (*C. sinensis* X *C. paradisi*), neuf tangors (*C. reticulata* x *C. sinensis*), deux pomelos (*C. paradisi*), une bigarade (*C. aurantium*) et une orange douce (*C. sinensis*). Les structures phylogénomiques de l'oranger doux, du clémentinier, du bigaradier et du mandarinier 'Ponkan' ont été décrites par Wu *et al.* (2014) à partir de données de re-séquençage complet et vont permettre de valider notre méthode d'identification des SNPs diagnostiques des taxons de base et du déchiffrement phylogénomique des espèces secondaires et variétés modernes.

Tableau 4 : Liste des espèces et des variétés du genre *Citrus* analysées

Nom Commun	Nom Latin (Swingle & Reece, 1967)	Référence
Chandler	<i>Citrus maxima</i> (Burm.) Merr	SRA 608
Deep red	<i>Citrus maxima</i> (Burm.) Merr	SRA 757
KaoPan	<i>Citrus maxima</i> (Burm.) Merr	SRA 321
Pink	<i>Citrus maxima</i> (Burm.) Merr	SRA 322
Tahiti	<i>Citrus maxima</i> (Burm.) Merr	SRA 727
Timor	<i>Citrus maxima</i> (Burm.) Merr	SRA 707
Chios	<i>Citrus reticulata</i> Blanco	ICVN 0100598
Cleopatra	<i>Citrus reticulata</i> Blanco	ICVN 0110066
Dancy	<i>Citrus reticulata</i> Blanco	SRA 594
Fuzhu	<i>Citrus reticulata</i> Blanco	ICVN 0100775
King	<i>Citrus reticulata</i> Blanco	SRA 583
Ladu	<i>Citrus reticulata</i> Blanco	SRA 595
Ponkan	<i>Citrus reticulata</i> Blanco	SRA 584
San Hu Hong Chu	<i>Citrus reticulata</i> Blanco	SRA 769
Satsuma Owari	<i>Citrus reticulata</i> Blanco	SRA 221
Sunki	<i>Citrus reticulata</i> Blanco	ICVN 0110078
Szibat	<i>Citrus reticulata</i> Blanco	SRA 596
Carvalho	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 271
Clemenules clementine	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	ICVN 0100389
Dweet	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	ICVN 0100464
Ellendale	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 656
Kiyomi	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 704
Murcott	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	ICVN 0100601
Ortanique	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 110
Temple	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 280
Allspice	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 327
Mapo	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 450
Minneola	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 451
Orlando	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 46
Pearl	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 296
Sampson	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	ICVN 0100456
San Jacinto	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 297
Seminole	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 298
Sunrise	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 457
Sunshine	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	ICVN 0100458
UGLI®	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 950
Webber SG	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	ICVN 0100475

Tableau 4 : Liste des espèces et des variétés du genre *Citrus* analysées (Suite)

Nom Commun	Nom Latin (Swingle & Reece, 1967)	Référence
Fairchild	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 30
Nova	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 158
Osceola	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 48
Page	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 159
Robinson	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 47
Fallglo	<i>Citrus reticulata</i> Blanco	IVIA 466
Kara	<i>Citrus reticulata</i> Blanco	SRA241
Wilking	<i>Citrus reticulata</i> Blanco	SRA112
Jackson	<i>Citrus sinensis</i> (L.) Osbeck x <i>Citrus paradisi</i> Macfad.	SRA 472
Triumph	<i>Citrus sinensis</i> (L.) Osbeck x <i>Citrus paradisi</i> Macfad.	ICVN 0100474
Wekiwa SG	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 462
Fortune	<i>Citrus reticulata</i> Blanco	SRA 31
Fremont	<i>Citrus reticulata</i> Blanco	SRA 147
Ambersweet	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 875
Marsh	<i>Citrus paradisi</i> Macfad.	SRA 188
Star Ruby	<i>Citrus paradisi</i> Macfad.	ICVN 0100293
Bigaradier Maroc	<i>Citrus aurantium</i> L.	ICVN 0110033
Washington Navel	<i>Citrus sinensis</i> (L.) Osbeck	SRA 203

2. Etude du polymorphisme moléculaire et méthodes utilisées

2.1 Amplification de l'espaceur intergénique *trnL*_(UAA)-*trnF*_(GAA)

La région non codante *trnL*_(UAA)-*trnF*_(GAA) de l'ADN chloroplastique a été amplifiée en utilisant deux amorces sens et anti-sens e : 5'-GGTTCAAGTCCCTCTATCCC-3' et f : 5'-ATTGAACTGGTGACACGAG-3', du gène *trnL*_(UAA) comme a été décrit par Tabarlet *et al.* (1991) (Figure 26).

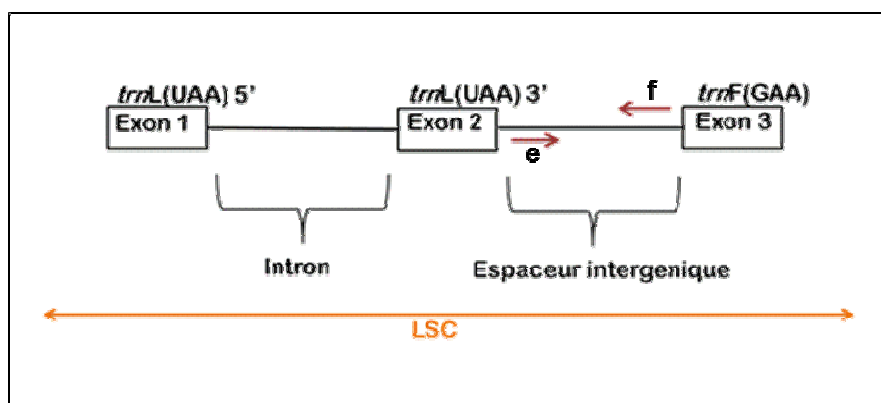


Figure 26 : Structure des sites des amorces désignées par Tabarlet *et al.* (1991) pour l'amplification de l'espaceur intergénique *trnL*_(UAA)-*trnF*_(GAA)

2.2 Le séquençage selon la technique de Sanger

Le séquençage des amplimères de l'espaceur intergénique *trn*_{L_{UAA}}-*trn*_{F_{GAA}} des échantillons tunisiens ont été effectué par la méthode automatisée de séquençage fluorescent; Genetic Analyser 3130 (Applied Biosystems, Foster City CA, USA). Le séquençage consiste à déterminer l'enchainement linéaire des nucléotides d'un génome ou d'un fragment d'ADN. La technique consiste à synthétiser un brin d'ADN complémentaire d'un brin matrice par l'ADN polymérase en présence de désoxyribonucléotides triphosphates (dNTPs). Des didésoxyribonucléotides triphosphates (ddNTPs) marqués par fluorescence ont été ajoutés au mélange servant de terminateurs d'élongation de façon aléatoire dans la réaction. La lecture de la séquence se fait en faisant migrer les fragments obtenus sur gel polyacrylamide. Le séquençage Sanger a été par la suite automatisé. Ainsi, les machines utilisant des gels de polyacrylamide ont rapidement été supplantées par les séquenceurs capillaires. Cette technologie principalement représentée par le 3730 DNA Analyser de Applied Biosystems, depuis sa commercialisation en 2002, est maintenant considérée comme la première génération du séquençage haut-débit.

2.3 Génotypage par la méthode KASPar™

Le génotypage de nos variétés a été réalisé en utilisant la méthode KASPar™ développée par KBioscience. Il s'agit d'un système de génotypage SNP par la technique PCR compétitive allèle spécifique. Elle est basée sur le système double FRET (*Fluorescent Resonance Energy Transfert*) et permet ainsi la détection des mutations ponctuelles (SNP). Le principe de la méthode repose sur l'utilisation de deux cassettes FRET fluorescentes, HEX ou FAM, qui sont conjuguées à l'amorce mais désactivées via le transfert de l'énergie de résonance. L'ADN est amplifié avec un thermocycleur utilisant des amorces allèles spécifiques, entraînant la séparation du fluorochrome et de la Quencher quand l'amorce de la cassette FRET est hybridée avec l'ADN (Cuppen, 2007) (Figure 27). Les amorces ont été conçues par LGC Genomics® basées sur les séquences flanquant le locus SNP d'approximativement 50 nucléotides de chaque côté (Annexe 3). Les signaux fluorescents du produit PCR ont été mesurés avec Fluostar Omega (BMG) et l'appel des génotypes a été effectué avec le logiciel KlusterCaller (LGC Genomics).

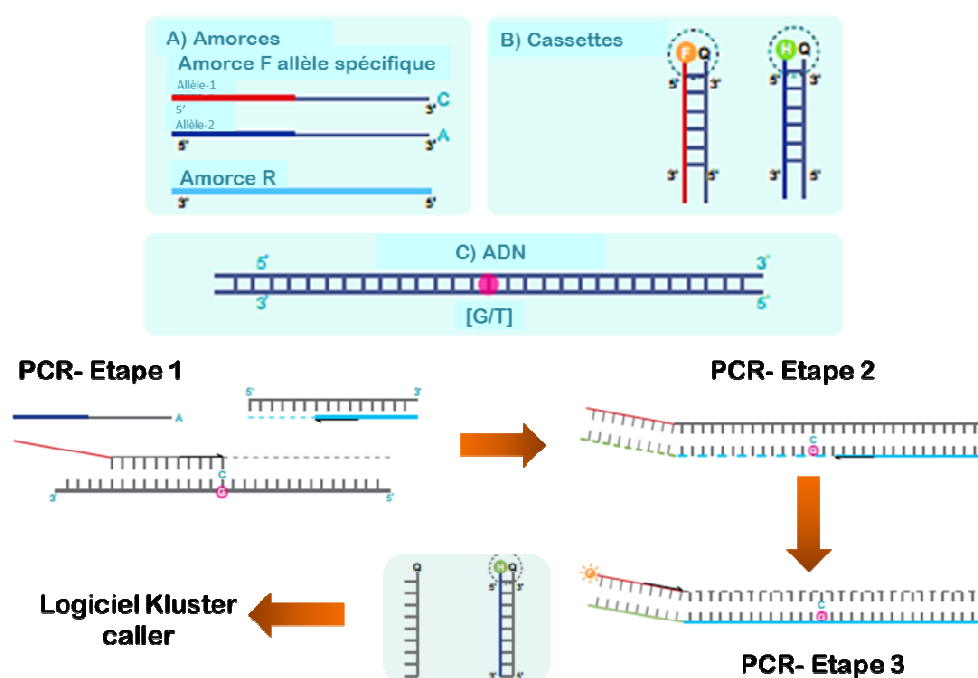


Figure 27 : Les étapes du génotypage par la méthode KASPar™ (Images de KBioscience, assemblage et montage par Amel Oueslati Bahri)

2.4 La technique de génotypage par séquençage (GBS : Genotyping By Sequencing)

La technique GBS est une méthode de séquençage à haut débit basée sur la construction de bibliothèques permettant de réduire la complexité du génome après une restriction enzymatique (Poland & Rife, 2012). Cette technique, permet de séquencer ou de re-séquencer en parallèle dans un même *run* un nombre important d'individu tout en conservant une profondeur suffisante pour l'appel des génotypes (SNPs et InDels ; Elshire *et al.*, 2011).

2.4.1 Préparation des bibliothèques et séquençage

La concentration de l'ADN génomique a été ajustée à 20 ng/μL. Les bibliothèques multiplexes GBS ont été préparées selon le protocole décrit par Elshire *et al.* (2011). Chaque échantillon d'ADN a été digéré par une enzyme de restriction *ApeKI* (New England Biolabs, Hitchin, Royaume-Uni). Le produit a été ligué en utilisant une ligase enzymatique *T4 DNA* (New England Biolabs, Hitchin, Royaume-Uni) à 22°C pendant 1 h. Une PCR a été ensuite élaborée durant laquelle une base a été ajoutée (A) après le motif de l'enzyme de restriction afin d'augmenter la profondeur de lecture en réduisant le nombre des sites amplifiés (Sonah

et al., 2013). Les librairies ont ensuite été séquencées avec un *Illumina HiSeq 2000* à la plate-forme MGX à Montpellier, France.

2.4.2 L'appel des génotypes SNPs et InDels

Un traitement des données brutes issues du séquençage *Illumina HiSeq 2000* de banques GBS a été effectué à l'aide du pipeline informatique (*Tassel 4.0 GBS pipeline*) (Glaubitz et al., 2014), avec un ancrage des séquences obtenues sur la séquence du génome de référence de l'haploïde de clémentinier établi par l'ICGC (Wu et al., 2014 ; <https://www.citrusgenomedb.org/species/clementina/genome1.0>).

Le programme repère les lectures (*reads*) correspondant à une même région, repère les polymorphismes (SNP et InDel) et positionne les séquences ainsi repérées sur la séquence de référence du clémentinier en utilisant Bowtie2 v2.2.6 (Langmead & Salzberg, 2012). Les sites SNPs et InDels pour lesquels il y avait 50% de données manquantes ont été éliminés. Un premier filtre a été performé au niveau des loci.

La procédure d'identification des marqueurs diagnostiques de la différenciation *C. maxima* /*C. reticulata* est décrite dans la partie des résultats et discussion. Un filtre additionnel a été effectué pour supprimer et remplacer par « données manquantes : N », les génotypes qui avaient moins de 5 lectures pour le site considéré et l'individu considéré (Petr Danecek et al., 2011).

3. Analyse *IN SILICO* des séquences de l'espaceur intergénique *trnL_(UAA)-trnF_(GAA)*

3.1 Vérification de l'identité des séquences

Une première étape des analyses des séquences de l'espaceur intergénique *trnL_(UAA)-trnF_(GAA)* consiste à la collecte de séquences homologues au sein des banques de données. L'outil couramment utilisé est le programme d'outil de recherche d'alignement local élémentaire (BLAST : *Basic Local Alignment Search Tool*). Il s'agit d'une méthode heuristique permettant de rechercher les régions similaires entre deux ou plusieurs séquences de nucléotides. Cette méthode permet d'obtenir à partir d'une séquence donnée une centaine de séquences homologues potentiellement intéressantes. Ce logiciel est disponible en ligne sur le site du serveur NCBI à l'adresse suivante : <http://www.ncbi.nlm.nih.gov/Blast>.

Le logiciel BLAST (Altschul et al., 1997) calcule ainsi l'homologie entre une séquence donnée et une banque de séquence en tenant compte du coefficient d'homologie E (E-value) et du score. Les alignements obtenus sont ensuite triés par ordre croissant de la E-value (qui représente le nombre de résultats attendus par le hasard). Ainsi, une «E-value» faible

(proche du 0) indique qu'un résultat est statistiquement significatif (ou fiable). Egalement, le choix de la matrice de substitution est un paramètre important de la méthode BLAST. Par défaut, les programmes BLAST implémentés sur le serveur NCBI proposent la matrice BLOSUM62. Ainsi, ce filtrage réduit le nombre des séquences tout en préservant l'information structurale pertinente.

3.2 Visualisation, correction et alignement des séquences

L'édition, l'alignement et l'analyse des séquences de l'espaceur intergénique $trnL_{UAA}$ - $trnF_{GAA}$ ont été réalisés avec le programme BioEdit version 6.0.5. (Hall, 1999). Ce logiciel permet la visualisation des séquences nucléotidiques. Les séquences, obtenues depuis le séquenceur, sont représentées sous forme de chromatogrammes indiquant chaque nucléotide par une couleur différente. La correction manuelle, en utilisant la fonction Edit du logiciel BioEdit, des séquences consiste à vérifier la correspondance entre le chromatogramme obtenu et sa base attribuée.

Les séquences d'ADN obtenues après la correction manuelle ont été alignées par l'algorithme ClustalW du logiciel BioEdit version 6.0.5. (Paramétrages: *options full multiple alignment*). L'alignement optimal est basé sur un critère mathématique qui est défini à partir d'une matrice de similarité. Ainsi, l'alignement est destiné à découvrir les séquences homologues qui selon l'hypothèse, héritées d'un ancêtre commun, ont divergé indépendamment après la séparation des espèces. Une fois alignées, la comparaison des séquences homologues permet donc de repérer des différences au niveau de certains sites entre les différentes séquences alignées. Les changements ainsi observés représentent les sites variables qui sont dus à un changement au niveau d'une base. Des « gaps » peuvent être aussi identifiés au niveau d'un alignement ; ils correspondent soit à une insertion soit à une délétion d'une base. L'ensemble de ces changements correspondent à des mutations. L'alignement des séquences a été sauvegardé au format fasta pour pouvoir être transféré vers les logiciels de phylogénie MEGA version 6 (Tamura *et al.*, 2013).

3.3 Identification et annotation des séquences du pseudogène $trnF_{(GAA)}$

Au niveau du site Repbase (<http://www.girinst.org/>), nous avons identifié le pseudogène $trnF_{(GAA)}$ par l'outil CENSOR en utilisant les options « *mask pseudogenes* » et « *report simple repeats* » (Kohany *et al.*, 2006). Les séquences ainsi identifiées ont été comparées à des séquences d'eucaryotes connus et des prototypes stockés au niveau de la base de données Repbase en utilisant le logiciel WU-BLAST.

Après l'identification des pseudogènes au niveau de l'espaceur intergénique $trnL_{UAA}$ - $trnF_{GAA}$, les éléments structuraux du gène $trnF_{(GAA)}$ ont été annotés manuellement en comparant avec le génome chloroplastique complet de *Solanum bulbocastanum* Dunal

(DQ347958). Afin d'identifier les répétitions pseudogéniques et déterminer les sites des enzymes de restriction, nous avons eu recours au programme Geneious v.4.8.5 (Biomatters Ltd).

3.4 Taille, composition nucléotidique et identification des différents types de mutations au niveau des séquences étudiées

La détermination du pourcentage en GC et AT ainsi que l'étude de la fréquence nucléotidique de la région étudiée ont été élaborés en utilisant le logiciel MEGA 6 (Tamura *et al.*, 2013). Ceci nous a permis d'estimer la composition nucléotidique et d'évaluer la complexité de l'intergène *trnL_{UAA}-trnF_{GAA}*. Une comparaison de la taille des séquences a été effectuée afin d'évaluer la variation de séquence au sein et entre les espèces étudiées. Les différents modèles substitutionnels nucléotidiques ont aussi été déterminés par le logiciel MEGA 6 afin de cibler les événements mutationnels établis au cours de l'évolution. Ainsi le taux de transition/transversion et le rapport R de la totalité des bases azotées ont été également calculés.

$$R = \frac{[A * G * K1 + T * C * K2]}{[(A + G) * (T + C)]}$$

Avec :

K1 : le taux de transition/transversion pour les bases puriques

K2 : le taux de transition/transversion pour les bases pyrimidiques

A, G, C et T : les fréquences nucléotidiques

4. Identification et sélection des SNPs

Les séquences de chaque région chloroplastique ont été alignées avec la séquence de référence du génome chloroplastique *C. sinensis* (Bausher *et al.*, 2006) disponible dans la base de données « GenBank » (NC008334) et ce, en utilisant le logiciel BioEdit (Hall, 1999). Par la suite, les séquences ont été classées selon l'arbre consensus généré par Bayer *et al.* (2009) afin de sélectionner les SNPs entre les clades aux différents niveaux taxonomiques. Un SNP considéré comme diagnostique pour un clade donné lorsque toutes les accessions au sein de ce clade sont identiques et différentes de toutes les accessions qui n'appartiennent pas à ce clade.

Les loci SNPs ont, en outre, été sélectionnés pour le développement de marqueurs KASPar selon des critères généraux de conception de marqueurs permettant ultérieurement leur utilisation chimique. En effet, ces critères de conception exigent la disponibilité d'une séquence de 50pb en amont et en aval du site SNP avec 15pb autour du SNP sélectionné sans polymorphisme moléculaire additionnel.

5. Paramètres génétiques

L'étude du polymorphisme moléculaire se base sur différents paramètres statistiques qui ont été estimés, lors de cette étude, en ayant recours au logiciel DnaSP version 5.10.01 (Librado & Rozas, 2009).

5.1 Diversité haplotypique

La diversité haplotypique (H_d) ou la diversité génique est l'équivalent de l'hétérozygotie attendue pour les données diploïdes. C'est la probabilité que deux haplotypes tirés au hasard dans un échantillon soient différents (Nei, 1987).

$$H_d = \frac{n}{n-1} \left(1 - \sum_{i=1}^k p_i^2 \right)$$

Avec :

n : nombre de séquence de l'échantillon

k : nombre d'haplotypes différents

p_i : fréquence de l'haplotype i dans l'échantillon

5.2 Diversité nucléotidique

La diversité nucléotidique P_i est la probabilité que deux nucléotides homologues choisis au hasard soient différents (Nei, 1987 ; Tajima, 1983).

$$P_i = \frac{\sum_{i=1}^k \sum_{j < i} p_i p_j d_{ij}}{L}$$

Avec :

d_{ij} : une estimation du nombre de mutations apparues depuis la divergence des haplotypes i et j

k : le nombre d'haplotypes

p : la fréquence de l'haplotype

L : nombre des loci

Plusieurs paramètres génétiques ont été calculés via le logiciel Excel, qui sont : le taux d'hétérozygotie, l'indice de fixation et l'indice de différenciation génétique

5.3 Le contenu informatif du polymorphisme (PIC)

C'est le nombre moyen de marqueurs polymorphes par locus, variant de 0 à 1. La valeur moyenne du PIC met en évidence les loci les plus polymorphes sur l'ensemble des individus étudiés. Ce paramètre a été calculé via le logiciel Excel. Il est estimé pour chaque marqueur selon la formule ci-dessous (Botstein *et al.*, 1980) :

$$PIC = 1 - \sum_{i=1}^k p_i^2$$

k: le nombre total

p_i : la fréquence du $i^{\text{ème}}$ allèle pour tous les génotypes obtenus

5.4 La fréquence allélique F

La fréquence allélique F correspond au nombre moyen d'allèle par locus. Cependant, cet estimateur est étroitement lié à la taille de l'échantillon ; plus le nombre d'individus est grand plus le nombre d'allèles observés sera grand. Ce paramètre a été calculé via le logiciel Excel.

5.5 Le taux d'hétérozygotie observé et attendue

L'hétérozygotie mesure le taux d'hétérozygotes dans une population donnée. C'est un paramètre largement utilisée dans les études de la diversité génétique à l'échelle de la population ou de l'individu. Ce paramètre a été calculé via le logiciel Excel. Il peut être définie de deux manières différentes (Kaeuffer, 2008).

L'hétérozygotie observée H_o (indice de diversité de Nei, 1973) correspond, pour un locus donné, à l'ensemble des individus hétérozygotes observés dans l'échantillon divisé par le nombre total des individus étudiés.

L'hétérozygotie attendue H_e , sous l'hypothèse de Hardy-Weinberg dans le cas d'une population panmictique, correspond à la probabilité de tirer simultanément au hasard deux individus différents. Il peut être calculé pour un locus donné de la façon suivante (Nei, 1978):

$$H_e = 1 - \sum_{i=1}^n p_i^2$$

Où p_i est la fréquence du $i^{\text{ème}}$ allèle au locus considéré

5.6 Indice de fixation F_{IS} de Wright

Le paramètre F_{IS} de Wright, dénommé aussi indice de fixation et appelé auparavant coefficient de consanguinité (Wright, 1969). La répartition de la diversité ou la structuration des populations sont généralement estimées par l'indice F_{IS} calculé selon la formule suivante via le logiciel Excel (Wright, 1969 ; Weir & Cockerham 1984):

$$FIS = \frac{H_e - H_o}{H_e} = 1 - \left(\frac{H_o}{H_e} \right)$$

Avec

H_o : l'hétérozygotie observée

H_e : l'hétérozygotie attendue, calculée à partir des fréquences alléliques dans l'hypothèse de Hardy-Weinberg

L'indice de fixation F_{IS} reflète la différenciation de l'hétérozygotie des individus à l'intérieur des populations. Ainsi, une valeur de F_{IS} positive reflète un défaut d'hétérozygotie dans la population qui peut être due à la consanguinité ou à des flux de gènes limités entre sous populations de la population globale. Si le F_{IS} est inférieur à 1 cela est expliqué par une hétérozygotie excédentaire. Quand le F_{IS} est nul, on parle d'une population à équilibre de Hardy-Weinberg.

5.7 Analyse de la différenciation génétique

5.7.1 Indice de différenciation génétique F_{ST}

Wright (1965 ; 1978) a défini l'indice F_{ST} comme l'hétérogénéité des fréquences alléliques entre subdivisions d'une population. Il représente la corrélation entre allèles à l'intérieur d'une sous-population par rapport à l'ensemble des sous populations. Cet indice permet de quantifier le niveau de différenciation entre groupes en fonction des paramètres F_{IS} et F_{IT} .

$$F_{ST} = 1 - \frac{1 - F_{IT}}{1 - F_{IS}}$$

Avec :

F_{IS} : différenciation des individus à l'intérieur des populations

F_{IT} : différenciation des individus par rapport au total

Ainsi, d'après Wright (1978), une valeur du F_{ST} comprise entre 0 et 0,05 indique une faible différenciation, une valeur de F_{ST} comprise entre 0,05 et 0,15 indique une différenciation modérée et une valeur de F_{ST} comprise entre 0,15 et 0,25 suggère une importante différenciation. Quand la valeur du F_{ST} est supérieure à 0,25 elle est révélatrice d'une très importante différenciation.

5.7.2 Indice de différenciation génétique G_{ST}

L'indice G_{ST} de Nei est une redéfinition du F_{ST} de Wright. Il s'agit d'une mesure de la différenciation génétique entre sous-populations aux loci multialléliques (Nei, 1977) :

$$G_{ST} = 1 - \frac{H_S}{H_T}$$

Avec :

H_S : la moyenne des diversités génétiques intra-populations

H_T : la diversité génétique totale

L'indice G_{ST} a été calculé via le logiciel DNAsP au cours de l'analyse de l'espaceur intergénique *trnL-trnF*.

5.7.2.1 Méthode utilisée pour l'analyse des données KASPar

Au cours de cette analyse, le paramètre G_{ST} (Nei, 1973) s'écrit en fonction des fréquences alléliques. Ce paramètre a été utilisé afin d'estimer l'efficacité de chaque marqueur développé pour différencier le clade précédemment identifié calculé via le logiciel Excel. L'évaluation du G_{ST} a été effectuée en utilisant le logiciel Excel en considérant deux sous-populations. La première concerne les variétés théoriquement incluses dans le clade concerné (Ci) et la seconde inclue la population théorique de toutes les autres variétés (C-i). Les analyses ont été réalisées à partir des fréquences alléliques estimées pour chaque taxon en considérant la même taille de population afin d'estimer la fréquence des deux sous-populations (Ci et C-i) et la fréquence de la population totale (Tot).

$$GST_{cladei} = \frac{He_{Tot} - \frac{He_{Ci} + He_{C-i}}{2}}{He_{Tot}}$$

He représente la diversité génétique intra-population calculée selon la formule suivante :

$$He = 1 - \sum_{i=1}^n p_i^2$$

Où p_i est la fréquence du $i^{ème}$ allèle dans la population ou la sous-population considérées

5.7.2.2 Méthode utilisée pour l'analyse des données GBS

La recherche des SNPs diagnostiques de la différenciation de *C. maxima* et de *C. reticulata* était basée sur l'estimation du paramètre G_{ST} (Nei, 1973). Cet indice de différenciation génétique a été calculé, via le logiciel Excel, à partir de l'estimation de la fréquence allélique de chaque taxon de base en considérant une même taille populationnelle pour chaque taxon afin d'estimer la fréquence de la population totale selon la formule suivante.

$$GST = \frac{He_{Tot} - \frac{(He_{reticulata} + He_{maxima})}{2}}{He_{Tot}}$$

Les valeurs du G_{ST} varient de 0 à 1. Une faible valeur indique un faible niveau de différenciation entre les sous-populations, alors qu'une valeur proche de 1 indique qu'une part majeur de la variabilité réside entre les sous-populations. Une valeur du G_{ST} égal à 1 signifie que les deux taxons sont totalement différenciés pour les marqueurs étudiés.

5.7.3 Le coefficient de différenciation génétique NST

Le coefficient de différenciation génétique NST prend en compte la similarité entre les haplotypes. Il permet de tester l'impact de la mutation sur la différenciation génétique entre populations en mesurant les distances génétiques entre les haplotypes (Pons & Petit, 1996).

5.7.4 Flux de gènes et nombre de migrants efficaces Nm

Selon Slatkin et Barton (1989), la quantification des flux de gènes entre populations se fait dans un modèle théorique dit 'en île', par le calcul du nombre de migrants efficaces par génération (Nm) qui peut être mesuré à l'aide du paramètre G_{ST} . Ainsi, Plus la valeur de Nm est supérieure à 1, plus l'échange de gènes est important.

$$Nm = \frac{1 - G_{ST}}{4 G_{ST} \left(\frac{n}{n-1} \right)^2}$$

Avec :

n : le nombre de populations

5.8 Analyse cladistique

La phylogénie permet la comparaison et le rapprochement des séquences en se basant à la fois sur des paramètres purement mathématiques et statistiques et sur des données biologiques d'évolution. Un arbre phylogénétique construit en se basant sur la méthode phénétique est un arbre qui repose sur le postulat de base qu'il y a corrélation entre le degré de ressemblance et le degré de parenté des taxons étudiés. Il s'agit ici, d'une ressemblance globale établie à partir du maximum d'observations disponibles constituant un ensemble de nature homogène (exemple : ensemble de fréquences alléliques, séquence nucléotidique d'un même gène...).

5.8.1 La méthode du plus proche voisin (NJ : *Neighbor-joining*)

Les comparaisons phylogénétiques des séquences de même taille ont été réalisées à partir du logiciel DARwin (Perrier & Jacquemoud Collet, 2006). Ces méthodes tentent de quantifier la ressemblance entre les organismes en calculant un indice de similitude globale entre deux taxons. Cet indice exprimé en pourcentage est assimilé à la distance génétique entre les deux taxons. Dans le cas des séquences alignées, cette distance est le nombre de nucléotides différents entre les deux espèces, divisé par le nombre de sites examinés (Lecointre *et al.*, 2001). L'analyse des séquences par cette méthode se fait en premier lieu en calculant les distances et en second lieu en construisant l'arbre phylogénétique.

Ainsi, la matrice des caractères (alignement) doit être transformée en matrice de distances évolutives séparant deux individus homologues (Darlu & Tassy, 1993). Ce modèle évolutif décrit de manière statistique le processus évolutif agissant sur la séquence (Jukes-

Cantor, Tajima-Nei, Kimura). La notion de distance découle de celle de similitude: plus la similitude entre deux individus est forte plus la distance entre eux est faible. Une fois les distances calculées, il existe plusieurs méthodes permettant de construire l'arbre phylogénétique à partir de la matrice des distances.

En se basant sur la matrice des distances, un arbre de classification de type Neighbour-Joining a été dressé par l'outil du logiciel DARwin. La robustesse des nœuds a été évaluée à partir de 10 000 bootstraps. Le Neighbour-Joining est la méthode la plus préconisée actuellement, c'est une méthode heuristique basée sur le principe de minimum d'évolution qui permettra de discriminer entre les différentes variétés et de définir les chlorotypes présents (Saitou & Nei, 1987).

Les méthodes de distances sont rapides et permettent l'analyse de grandes bases de données tout en testant un grand nombre d'hypothèses alternatives. Elles permettent également d'intégrer des modèles de changements évolutifs non intégrables dans d'autres méthodes.

De même L'analyse Neighbor-joining a été utilisée pour analyser la différenciation entre les variétés représentatives des taxons de bases *C. maxima* et *C. reticulata*. Le principe de la méthode a été décrit précédemment. Sauf que dans cette étude, l'arbre NJ s'est basé sur une matrice de dissimilarité génétique en utilisant l'indice « simple matching» (d_{i-j}) entre chaque paire de variétés. Il est calculé selon la formule suivante :

$$d_{i-j} = 1 - 1/L \sum_{l=1}^L m_l / 2$$

Avec

d_{i-j} : dissimilarité entre les unités i et j

L : nombre de loci

m_l : nombre des allèles assorti pour le locus l

5.8.2 La méthode de regroupement non pondéré par paire avec moyenne arithmétique : UPGMA

La méthode de regroupement non pondéré par paire avec moyenne arithmétique (UPGMA : *Unweighted Pair Group Method with the Arithmetic Average*) est une analyse cladistique qui permet de construire un arbre de classification des variétés à partir d'une matrice de distances génétiques. Elle permet de regrouper les individus les plus proches pour révéler les relations entre eux. La méthode UPGMA se base sur la valeur moyenne de la variance entre groupes d'individus qu'elle compare sur la base des distances estimées pour un ensemble de données quantitatives, qualitatives ou éventuellement la combinaison des deux types de variables. Le regroupement hiérarchique produit un arbre de classification appelé un dendrogramme.

Cette analyse cladistique a été élaborée à partir des données SNPs KASPar à partir d'une matrice de distance Euclidienne effectuée entre chaque paire de variétés en ayant recours au logiciel DARWin (Perrier & Jacquemoud Collet, 2006). La génération du dendrogramme UPGMA a été effectué par le recours au logiciel Mega version 6 (Tamura *et al.*, 2013).

5.9 Analyse factorielle sur tableau de distances

L'Analyse Factorielle sur Tableau de Distances (AFTD) constitue la mise en œuvre la plus simple de l'analyse factorielle. Elle est utilisée pour traiter une matrice de distances calculées avec un indice original (d_{i-j}), choisi de manière adaptée à ses propres données. Il est cependant nécessaire pour que la distance obtenue soit une distance euclidienne ou rendue euclidienne par transformation (Hamon *et al.*, 1999).

L'indice d_{i-j} a été estimé pour l'analyse factorielle en utilisant le logiciel DARwin (Perrier et Jacquemoud Collet, 2006). L'indice de similarité utilisé dans l'identification des introgressions dans les variétés représentatives des taxons de base a été calculé comme suit :

$$S_{i-j} = 1 - d_{i-j}$$

Avec

d_{i-j} : dissimilarité entre les unités i et j

6. Etude de l'évolution moléculaire

La diversité génétique observée au sein d'une population peut être la conséquence de plusieurs facteurs tels que la sélection, la dérive génétique et la mutation. Ainsi, la sélection naturelle entraîne un changement du profil de la diversité génétique par rapport à celui attendu sous neutralité. Une série de tests de neutralité a été développée permettant de savoir si le polymorphisme observé est compatible avec un modèle d'évolution neutre et en cas de sélection naturelle identifier ses multiples signatures. Plusieurs paramètres ont été calculé à l'aide du programme DnaSP (version 5.0) (Rozas *et al.*, 2009).

6.1 Test de Tajima

Le test D de Tajima (1989) est utilisé pour détecter des effets démographiques (goulot d'étranglement, expansion de population) ou sélectifs en évaluant une déviation générale du spectre de fréquences alléliques par rapport à l'attendu sous des conditions neutres. Ce test va mesurer la différence entre deux estimateurs θ du taux de mutations dans la population comme le décrit la formule suivante :

$$D = \frac{\theta_{\pi} - \theta_s}{\sqrt{\text{Var}(\theta_{\pi} - \theta_s)}}$$

Avec :

θ_s : le nombre total de sites polymorphes

θ_{π} : la moyenne du nombre de différences entre paires de séquences

Sous neutralité, θ_s et θ_{π} devraient être équivalentes. Ainsi, dans une population de taille constante non soumise à la sélection, la valeur du D de Tajima est égale à 0. Les valeurs seront négatives sous sélection positive ou sélection négative dues à un excès de mutations rares, et elles seront positives sous sélection balancée à cause d'un excès d'allèles à fréquence intermédiaire.

6.2 Test de Fu & Li

D'autres tests mesurant également un excès ou un déficit de mutations rares incluent les statistiques D^* et F^* de Fu & Li (1993). Ces tests sont similaires au D de Tajima mais font la distinction entre les mutations anciennes et récentes. Ils s'intéressent ainsi au nombre de singletons observés dans la population. Ces statistiques D^* et F^* comparent un estimateur du taux de mutation basé sur le nombre de singletons avec θ_s et θ_{π} , respectivement. Comme pour le D de Tajima, la valeur attendue pour ces deux tests sous neutralité est de zéro. Les valeurs négatives et positives indiquent respectivement une sélection positive/négative et une sélection balancée.

6.3 Test de neutralité de Fu

Le test de neutralité F_s de Fu (1997) est proche de celui de Tajima, seulement, il est basé sur la structure haplotypique dans le but de voir si le nombre d'haplotypes correspondrait au nombre de site polymorphes. Ce test compare le nombre moyen de différences deux à deux (π) avec le nombre d'haplotypes (h) dans la population. Il permet de tester les hypothèses d'excès ainsi qu'au déficit du nombre d'haplotypes étant donné le nombre de mutations observées. Le test de neutralité de Fu tend à être négatif quand il y a un excès de mutations récentes, c'est-à-dire d'allèles rares. Ce test s'avère être plus puissant que le test D de Tajima pour détecter les expansions démographiques. En appelant la probabilité S la formule de calcul du F_s de Fu s'écrit comme suit :

$$F_s = \ln\left(\frac{S}{1-S}\right)$$

Avec :

S : la probabilité

Résultats & Discussions

Chapitre I : Etude du polymorphisme génétique, de la phylogénie et de l'évolution moléculaire de la région non codante *trnL*_(UAA)-*trnF*_(GAA) de l'ADN chloroplastique chez les espèces tunisiennes du genre *Citrus*.

Chapitre II : Elaboration d'une clé moléculaire taxonomique de la sous famille des Aurantioideae en utilisant les marqueurs moléculaires SNPs diagnostiques des principaux clades génotypés par la PCR compétitive allèle-spécifique.

Chapitre III : Révélation des structures en mosaïque interspécifiques *C. maxima*/*C. reticulata* au niveau des génomes des variétés modernes de mandariniers et des hybrides de mandariniers par la technique de Génotypage par Séquençage (GBS).

Extraction de l'ADN génomique

L'extraction de l'ADN génomique est une étape cruciale dans une étude moléculaire. La qualité et la pureté des acides nucléiques comptent parmi les facteurs les plus critiques pour les analyses ultérieures. Pour ce faire, le recours aux kits commerciaux QIAcube (Qiagen, 2003) et le Mini kit « DNeasy Plant » (Qiagen S.A.; Madride, Espagne) nous ont permis de réaliser rapidement et efficacement l'extraction de l'ADN génomique qui a été évaluée sur un profil électrophorétique (Figure 28) comme un ADN hautement purifié avec absence de trainées. Par ailleurs, la mesure de la quantité d'ADN à l'aide du Nanodrop (ND-1000 version 1) a montré un intervalle de concentration de 51,2 ng/μL à 116 ng/μL avec un rapport de : DO260/DO280 variant de 1,8 à 2 témoignant de la pureté des ADNs obtenus et confirmant qu'ils sont exempts de toutes contamination (protéines, ARN et sels). L'allure et l'intensité relative des bandes de migration sur gel d'agarose sont bien parallèles aux valeurs de concentration d'ADN extrait pour les différents échantillons.

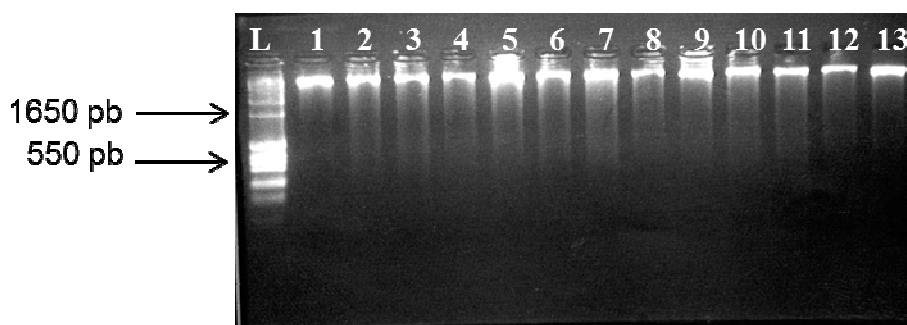


Figure 28 : Électrophorèse analytique sur gel d'agarose à 1% des ADNs génomiques totaux

- Piste L: Marqueur de taille (1Kb, Invitrogen)
- Piste 1→13: Extraits d'ADN génomiques totaux de quelques variétés d'agrumes selon le protocole QIAcube (Qiagen, 2003)

**Chapitre I : Etude du polymorphisme génétique,
de la phylogénie et de l'évolution moléculaire de
l'espaceur intergénique *trnL*_(UAA)-*trnF*_(GAA) de
l'ADN chloroplastique chez les espèces
tunisiennes du genre *Citrus***

1. Amplification de l'espaceur intergénique *trnL*_(UAA)-*trnF*_(GAA)

L'amplification de la région intergénique *trnL*_{UAA}-*trnF*_{GAA} a été appliquée sur 27 variétés tunisiennes en utilisant des amorces flanquant la région correspondante. Les résultats obtenus ont montré la présence d'une bande révélée chez tous les cultivars avec une taille d'environ de 400pb (Figure 29).

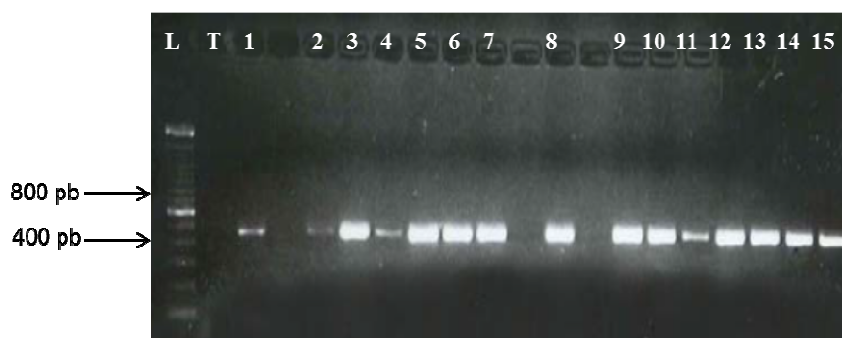


Figure 29 : Profil d'amplification de l'espaceur intergénique *trnL*-*trnF*

- Piste L: Marqueur de taille (100pb, Invitrogen)
- Piste T: Témoin négatif
- Piste 1: Produit PCR de la variété 'Minneola' de l'espèce *C. tangelo*
- Piste 2→7: Produit PCR de quelques quelques variétés de l'espèce *C. sinensis*
- Piste 8 : Produit PCR de la variété 'Bigaradier Maroc' de l'espèce *C. aurantium*
- Piste 9→12 : Produit PCR des variétés de l'espèce *C. reticulata*
- Piste 13→15 : Produit PCR de quelques quelques variétés de l'espèce *C. limon*

2. Séquençage de l'espaceur intergénique *trnL*_{UAA}-*trnF*_{GAA} de la région non codante chloroplastique

2.1 Vérification de l'identité des séquences par BLAST

Le séquençage Sanger des produits de PCR obtenus a permis de donner des chromatogrammes pour 27 différentes séquences appartenant à sept espèces du genre *Citrus*. Ces séquences ont été par la suite corrigées par l'outil du logiciel BioEdit (Figure 30). Un BLAST a été préalablement appliqué sur les séquences ainsi corrigées et a confirmé l'identité des séquences avec celles de la banque de données NCBI (Figure 31). Par ailleurs, une homologie absolue a été identifiée chez les séquences de l'intergène *trnL*_{UAA}-*trnF*_{GAA} des *Citrus* tunisiens en comparant avec plusieurs autres séquences de la même région des espèces du genre *Citrus*. Les séquences sont publiées dans la banque génomique GenBank avec une attribution de numéros d'accessions allant de KJ667680 à KJ667706 (Tableau 5).

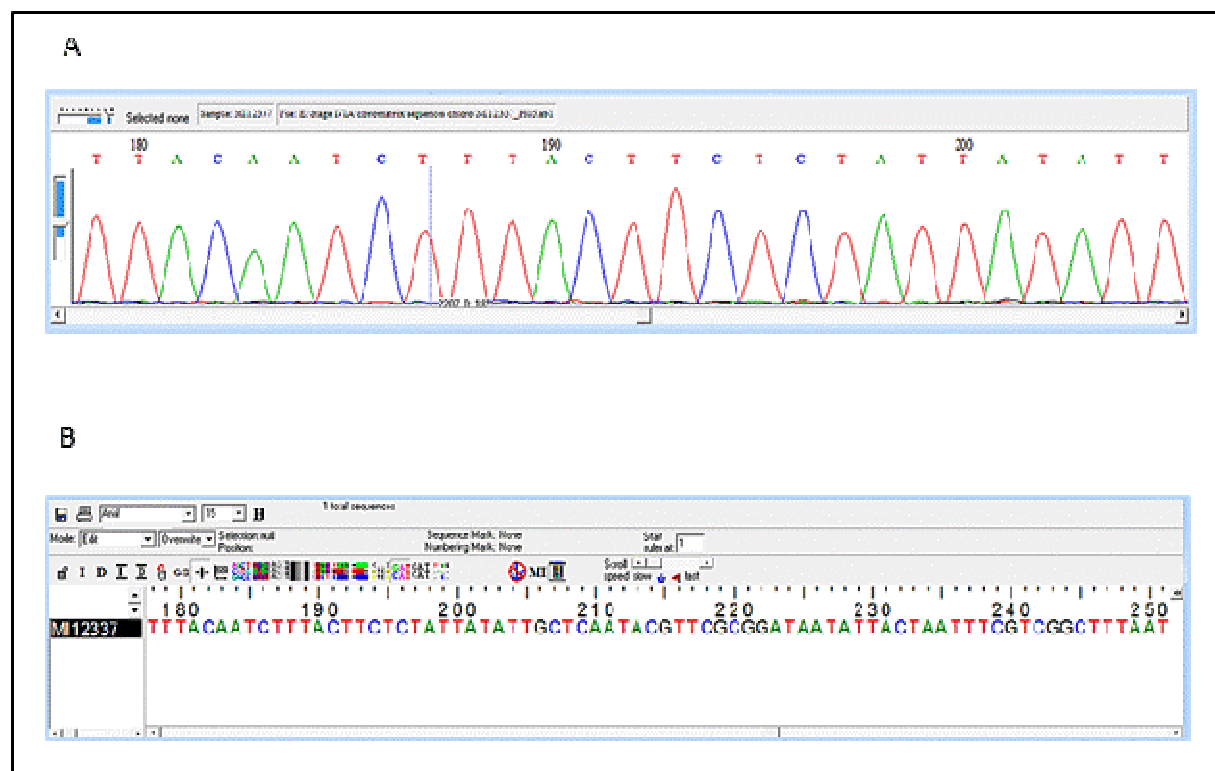


Figure 30 : Exemple de visualisation d'une séquence de l'espaceur intergénique *trnL_{UAA}-trnF_{GAA}* du cultivar 'Orange Sanguine' de l'espèce *Citrus sinensis* par le logiciel BioEdit

A : Chromatogramme ; B : Séquence d'ADN

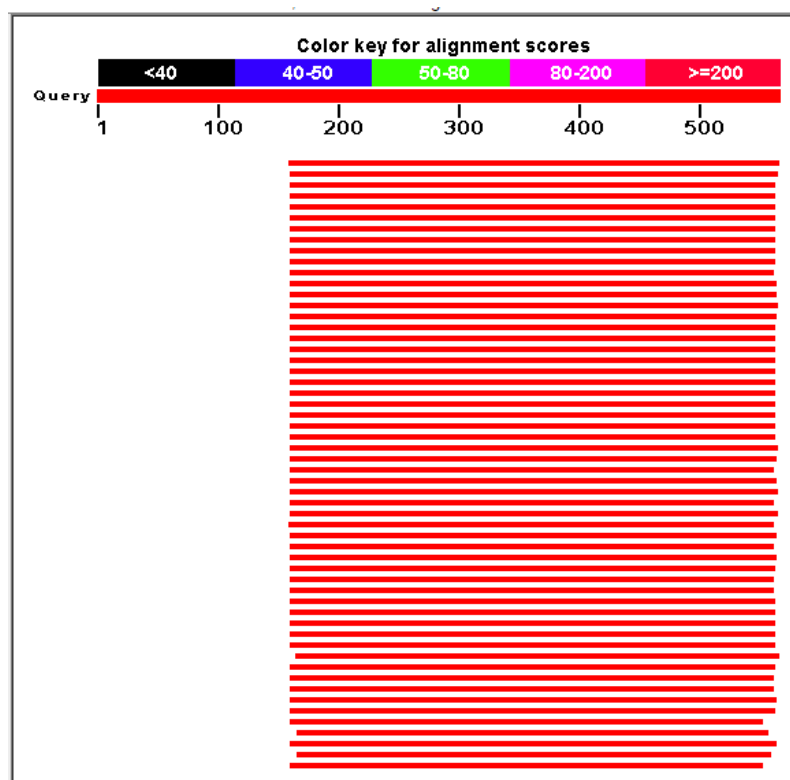


Figure 31 : Vérification de l'identité des séquences *trnL_{UAA}-trnF_{GAA}* par comparaison avec les séquences publiées dans la banque de données GenBank de NCBI

2.2 Analyse *in silico* des séquences de l'espaceur intergénique *trnL_{UAA}-trnF_{GAA}*

2.2.1 Variation de la longueur et de la composition nucléotidique des séquences *trnL_{UAA}-trnF_{GAA}*

La taille de l'intergène *trnL-trnF* varie de 331pb chez les variétés 'Encore' et 'Nova' de l'espèce *C. reticulata* à 415pb chez la variété 'Bigaradier du Maroc' de l'espèce *C. aurantium* avec une moyenne de 378,9pb (Tableau 5). Des longueurs de séquences similaires ont été identifiées au niveau de la même région de l'ADN chloroplastique de l'espèce *C. reticulata* (337pb) (Jena *et al.*, 2009). L'espaceur intergénique *trnL-trnF* a été utilisé dans plusieurs études moléculaires où les résultats ont montré une longueur moyenne de 460,9pb chez le figuier *Ficus carica* L. (Baraket *et al.*, 2009) et de 390,6pb chez le palmier dattier *Phoenix dactylifera* L. (Sakka *et al.*, 2013).

Tableau 5 : Caractéristiques des séquences des espèces du genre *Citrus*

Espèce	Nom commun	Numéro d'accension	%AT	%GC	Taille (pb)
<i>C. aurantium</i>	Bigaradier du Maroc Bigaradier	KJ667680	61,56	38,44	415
		KJ667681	61,56	38,44	411
<i>C. limon</i>	Feminello	KJ667682	61,56	38,44	410
	Monachello	KJ667683	61,56	38,44	410
	Santa Téréza	KJ667684	61,56	38,44	409
	Interdonato	KJ667685	61,56	38,44	412
<i>C. paradisi</i>	Pomelo Ruby	KJ667686	61,56	38,44	406
	Pomelo Marsh	KJ667687	61,56	38,44	407
	Pomelo Star Ruby	KJ667688	61,56	38,44	332
<i>C. sinensis</i>	Tarocco	KJ667689	61,56	38,44	406
	Maltaise barlerin	KJ667690	61,56	38,44	374
	Maltaise demi sanguine	KJ667691	61,56	38,44	386
	Navel Late	KJ667692	61,56	38,44	376
	Lane late	KJ667693	61,56	38,44	348
	New Hall	KJ667694	61,56	38,44	337
	Valencia Late	KJ667695	61,56	38,44	337
	Navelina	KJ667696	61,56	38,44	377
	Washington navel	KJ667697	61,56	38,44	341
<i>C. clementina</i>	Cassar	KJ667698	61,88	38,13	401
	Marisol	KJ667699	61,88	38,13	348
	Hernandina	KJ667700	61,88	38,13	402
	MA3	KJ667701	61,88	38,13	392
<i>C. tangelo</i>	Tangelo	KJ667702	61,56	38,44	391
<i>C. reticulata</i>	Ortanique	KJ667703	61,88	38,13	373
	Nova	KJ667704	61,88	38,13	331
	Fortune	KJ667705	61,88	38,13	367
	Encore	KJ667706	61,88	38,13	331
Moyenne	-	-	61,66	38,34	378,9
<i>C. maxima</i>		EF126663			
<i>C. medica</i>		EF126665			
<i>Atalantia Monophylla</i>		EF126636			

Les fréquences nucléotidiques sont de 30,35% (A), 31,31% (T), 20,00% (C), et 18,34% (G). Des résultats similaires ont été obtenus dans la région *trnL-trnF* chez *Citrus*, avec 29,9% de A, 32,3% de T, 19,6% de C et 18,3% de G (Jena *et al.*, 2009). Par ailleurs, le contenu en AT, d'une moyenne de 61,66%, est nettement supérieur au contenu en GC (une moyenne de 38,34%). Ceci confirme l'hypothèse prouvée par Shimada & Sugiura (1991) que chez plusieurs espèces végétales le pourcentage en GC est faible dans les espaceurs intergéniques, moyenne dans les introns et élevée dans les exons.

Les rapports de transition/transversion sont $k_1 = 0$ (purines) et $k_2 = 0$ (pyrimidines). Le rapport total de transition/transversion est $R = 0$. Nos résultats montrent l'absence des événements transitionnels et indiquent que les transversions sont les seules mutations substitutionnelles observées chez l'espaceur *trnL-trnF*. Cependant, les transversions de type T→A sont plus fréquents que celles de type G→T (Tableau 6).

Tableau 6 : Estimation du maximum de vraisemblance du modèle de la substitution nucléotidique au niveau de l'espaceur intergénique *trnL-trnF*

	A	T	C	G
A	-	15,65	10,00	0,00
T	15,17	-	0,00	9,17
C	15,17	0,00	-	9,17
G	0,00	15,65	10,00	-

Avec :

A : Adénine, T : Thymine, C : Cytosine, G : Guanine

Les substitutions transitionnelles sont en caractère **gras** et les substitutions transversionnelles sont en *italique*

L'alignement final de l'espaceur intergénique *trnL-trnF* s'étale sur 320pb sans gaps et données manquantes avec deux sites polymorphes qui sont d'une parcimonie informative (Tableau 5). Cependant, des études chloroplastiques antérieurs chez *Citrus* ont détecté, après exclusion des gaps et des données manquantes, 25 sites polymorphes dont 5 sont informatifs (Jena *et al.*, 2009).

En utilisant le programme Genious nous avons détecté au niveau de l'espaceur intergénique *trnL-trnF* 4 sites d'enzymes de restriction, qui sont : *PvuI* à la position 359, *AluI* aux positions 354 et 359, *MseI* aux positions 179, 272 et 290 et enfin l'enzyme de restriction *HinfI* a été localisé au niveau des positions 168, 184 et 239.

2.2.2 Diversité génétique

Une faible variation génétique a été enregistrée chez les espèces du genre *Citrus*. En effet, l'espaceur intergénique *trnL-trnF* nous a permis d'identifier 3 haplotypes représentant les 27 séquences des variétés du genre *Citrus* avec une diversité haplotypique égale à 0,476. La diversité nucléotidique (P_i) de l'espaceur *trnL-trnF* est faible et égale à 0,00199.

La theta (par site) est de 0,00162 et la moyenne des différences par paires de nucléotide (k) est de 0,638. L'espaceur intergénique a montré une absence des événements de recombinaison (Tableau 7). Des études antérieures ont montré une corrélation entre le taux de recombinaison et celui du polymorphisme. Ainsi, un taux faible ou une absence de recombinaison indique un très faible niveau de polymorphisme, ce qui confirme nos résultats (Kimura, 1983). D'autre part, la recombinaison ne joue pas un rôle majeur dans l'évolution des génomes cytoplasmiques dû à la prédominance de l'hérédité uniparentale et de l'haploïdie (Birky, 2001).

Tableau 7 : Paramètres du polymorphisme testés sur les séquences de l'espaceur intergénique *trnL-trnF*

Paramètres du polymorphisme	L'intergène <i>trnL-trnF</i>
Nombre des séquences	27
Taille de l'alignement (pb)	320
Site Monomorphe	318
Site Polymorphe (S)	2
Nombre des haplotypes (H)	3
Diversité Haplotypique (H_d)	0,476
Diversité Nucléotidique (P_i)	0,00199
Theta (par site) d'Eta	0,00162
Moyenne des différences par paires de nucléotide (K)	0,638
Nombre Minimum des événements de recombinaison (R_m)	0

2.2.3 Analyse phylogénétique de l'espaceur *trnL_{UAA}-trnF_{GAA}*

Les analyses phylogénétiques se basent sur 27 variétés tunisiennes et trois accessions extraites de la banque de données NCBI (*C. maxima*, *C. medica* et *Atalantia monophylla*). La topologie du dendrogramme NJ distingue entre trois groupes majeurs: (I), (II) et (III). Le premier groupe (I) comprend toutes les accessions appartenant aux espèces *C. reticulata* et *C. clementina*. Au sein de ce groupe se forment deux sous-groupes; le premier sous-groupe (I.1) comprend toutes les variétés appartenant à l'espèce *C. clementina* avec 'Ortanique' variété de l'espèce *C. reticulata*, cependant, le second sous-groupe (I.2) est constitué des variétés de l'espèce *C. reticulata*: 'Nova', 'Encore' et 'Fortune'. Le second groupe (II) est formé par une seule espèce *C. medica*, alors que, le troisième groupe (III) rassemble l'ensemble des variétés représentatives des espèces *C. tangelo*, *C. sinensis*, *C. limon*, *C. paradisi*, *C. aurantium* et *C. maxima* (Figure 32).

L'analyse phylogénétique des données des séquences chloroplastiques basées sur la région intergénique *trnL-trnF* montre une séparation entre les espèces *C. maxima*, *C. medica* et *C. reticulata* dans des groupes distincts ce qui supporte l'hypothèse de leur classification comme espèces de base de toutes les autres espèces du genre *Citrus*. Ce

concept a été soutenu par des études morphologiques, biochimiques et moléculaires (Scora 1975; Barrett & Rhodes 1976; Nicolosi *et al.*, 2000; Mabberley, 2004; Liang *et al.*, 2007; Pang *et al.*, 2007; Kumar *et al.*, 2012; Hynniewta *et al.*, 2014). Les espèces secondaires originaires des hybridations entre les espèces de base ont été prouvées par différentes études moléculaires (Nicolosi *et al.*, 2000; Barkley *et al.*, 2006; Garcia-Lor *et al.*, 2012; Ollitrault *et al.*, 2012).



Figure 32 : Un dendrogramme de type Neighbor-Joining des espèces tunisiennes du genre *Citrus* basé sur l'espaceur intergénique *trnL-trnF*. Les valeurs des Bootstrap (le pourcentage de bootstrap, >50%) résultant de 10 000 répliquions sont indiquées sur chaque branche

C. reticulata (la mandarine) est considérée comme l'une des espèces de base du genre *Citrus*. En effet, elle contribue dans le développement des fruits commerciaux des *Citrus* comme *C. clementina* (*C. reticulata* x *C. sinensis*), *C. tangelo* (*C. reticulata* x *C.*

paradisi), *C. sinensis* (backcross entre *C. reticulata* et *C. maxima*) et *C. aurantium* (*C. maxima* x *C. reticulata*). Des études antérieures ont montré d'une part que le groupe des mandarines inclut les mandarines comestibles comme *C. reticulata*, *C. clementina* et *C. unshiu* et d'autre part ils n'ont pas pu subdiviser ces espèces (Nicolosi *et al.*, 2000; Li *et al.*, 2007; Deng *et al.*, 2007; Urasaki *et al.*, 2005; Penjor *et al.*, 2013). Ainsi, l'arbre NJ construit à partir de nos données des séquences *trnL-trnF* a identifié clairement la présence de deux groupes *C. reticulata* et *C. clementina* (Figure 32).

C. clementina (la clémentine) s'agit d'un hybride interspécifique sélectionné avec un coup de chance par le père Clément en Algérie, à partir du plant *C. reticulata*. Cependant, *C. clementina* a été supposé être un hybride entre *C. reticulata* et *C. sinensis* (Deng *et al.*, 1996; Nicolosi *et al.*, 2000; Ollitrault *et al.*, 2012). L'hérédité maternelle de la clémentine à partir de la mandarine a été confirmée par des analyses du génome mitochondrial (Froelicher *et al.*, 2011). Par conséquent, le regroupement de la variété 'Ortanique' de l'espèce *C. reticulata* avec *C. clementina* supporte le lien de parenté entre ces deux espèces et indique que *C. reticulata* est le parent maternel de *C. clementina*.

D'autre part, nos données basées sur l'ADN chloroplastique montrent les espèces *C. sinensis*, *C. limon*, *C. paradisi*, *C. aurantium*, *C. tangelo* et *C. maxima* formant un groupe séparé ce qui est en cohérence avec les analyses phylogénétiques chloroplastiques de Bayer *et al.* (2009). Des études précédentes basées sur l'ADN chloroplastique ont montré que les espèces *C. aurantium*, *C. sinensis* et *C. limon* sont groupées à leur parent femelle *C. maxima* (Nicolosi *et al.*, 2000; Araujo *et al.*, 2003; Cheng *et al.*, 2005; Bayer *et al.*, 2009; Yamamoto *et al.*, 2011; Penjor *et al.*, 2013). En effet, nos résultats confirment la contribution du pool génétique *C. maxima* dans le développement de nos espèces secondaires.

C. aurantium (la bigarade) est une orange amère utilisée comme porte-greffe pour l'orange douce et d'autres biotypes de *Citrus* (Castle, 1987). Scora (1975) et Barrett & Rhodes (1976) ont montré que l'orange amère est un hybride entre *C. reticulata* et *C. maxima*. Cette hypothèse a été soutenue par une étude basée sur des isoenzymes (Fang, 1993). Plusieurs hypothèses ont été élaborées sur l'origine de l'espèce *C. aurantium*. Toute fois, Malik *et al.* (1974) a rejeté l'hypothèse de possibilité de croisement entre *C. sinensis* et *C. reticulata*. Les analyses moléculaires basées sur des données issues des marqueurs RAPD et SCAR ont révélés que le pamplemousse (*C. maxima*) a joué le rôle de parent femelle de l'orange amère (Nicolosi *et al.*, 2000). Cette hypothèse a été par la suite soutenue par Barkley *et al.* (2006) et Bayer *et al.* (2009). Li *et al.* (2010) a identifié que *C. aurantium* est un hybride entre *C. maxima* (parent maternel) et *C. reticulata* (parent paternel). Les données des séquences des génomes complets ont montré *C. maxima* parent maternel de *C. aurantium* (Carbonell-Cavallero *et al.*, 2015). Nos résultats montrent que *C. aurantium* ('Bigaradier' et 'Bigaradier du Maroc') se regroupent avec *C. tangelo*, *C. sinensis*, *C. limon*,

C. paradisi, *C. aurantium* et *C. maxima*. Ce résultat confirme que *C. maxima* est probablement le parent maternel de *C. aurantium*.

C. sinensis (l'orange douce) est le fruit d'un croisement entre *C. maxima* et *C. reticulata*. Ceci a été montré par Barrett & Rhodes (1976) dans une étude basée sur des marqueurs morphologiques et a été soutenue par une étude basée sur des marqueurs RAPD et SCAR (Nicolosi *et al.*, 2000). En plus, des données protéiques ont montré que *C. reticulata* est le parent de *C. sinensis* (Handa *et al.*, 1986; Malik *et al.*, 1974). D'autre part, des analyses chloroplastiques ont révélé que le génome de *C. sinensis* est dérivé de l'espèce *C. maxima* (Nicolosi *et al.*, 2000; Barkley *et al.*, 2006; Li *et al.*, 2010). Yingzhi (2007) a soutenu cette théorie et a montré un lien de parenté entre l'orange douce 'Valencia' et la mandarine et ce, en se basant sur des données de l'ADNr et a montré dans la même étude, en se basant sur des données de séquence de l'espaceur intergénique *trnL-trnF*, un lien de parenté avec le pamplemousse. Les marqueurs moléculaires de type AFLP et ITS ont montré que la mandarine est le parent paternel de *C. sinensis* (Li *et al.*, 2010). Quant aux analyses morphométriques et l'amplification dirigée d'ADN minisatellite (DAMD : *Directed Amplification of Minisatellite-region DNA*) ont prouvé que *C. maxima* est le parent putatif de l'orange douce (Kumar & Nair, 2013). Récemment, une étude a prouvé l'établissement d'un croisement entre le parent femelle *C. maxima* et la mandarine *C. reticulata*, l'hybride interspécifique ainsi obtenu a été croisé avec le parent mâle mandarine *C. reticulata* permettant la production de l'orange douce *C. sinensis* (Xu *et al.*, 2012) avec *C. maxima* parent femelle (Carbonell-Cavallero *et al.*, 2015). Selon nos données des séquences chloroplastiques, *C. sinensis* et *C. maxima* sont très proche génétiquement. Ces résultats indiquent l'origine hybride de *C. sinensis* impliquant *C. maxima* comme parent maternel ce qui supporte les résultats cités ci-dessus.

C. limon (Le citron) étant confirmé hybride demeure toutefois présenter une origine parentale ambiguë. L'utilisation des données moléculaires: suggère que *C. medica* et *C. aurantium* sont les parents de *C. limon* (Nicolosi *et al.*, 2000 ; Jena *et al.*, 2009 ; Li *et al.*, 2010). Les marqueurs moléculaires RAPD et SCAR et les données des séquences de l'ADNcp suggèrent *C. aurantium* parent maternel de *C. limon* (Nicolosi *et al.*, 2000 ; Li *et al.*, 2010). D'autres données moléculaires (dont AFLP et ITS) ont montré que *C. medica* est le parent paternel de *C. limon* (Federici *et al.*, 1998; Barkley *et al.*, 2006 ; Li *et al.*, 2010). Dans la présente étude, nos données des séquences de la région chloroplastique *trnL-trnF* montrent que *C. maxima*, *C. tangelo*, *C. sinensis*, *C. paradisi* et *C. aurantium* sont très proche de *C. limon*. Ces espèces sont impliquées directement ou indirectement comme ancêtres de *C. limon*. D'autre part, sur le NJ *C. medica* forme un groupe à part mais proche de *C. limon*. Ce qui exclu la possibilité de la contribution de *C. medica* comme parent maternel de *C. limon*. Toutes ces espèces sauf *C. maxima* possèdent *C. reticulata* comme

hybride ou ancêtre, ce qui explique la contribution génétique de la mandarine dans l'établissement de ces espèces secondaires. Ces résultats confirment les données ITS de Li *et al.* (2010).

C. paradisi (Le pomelo) est un hybride entre *C. maxima* et *C. sinensis* selon des données morphologiques (Barrett & Rhodes, 1976). Les données Isozymes (Fang, 1993; Herrero *et al.*, 1996) et moléculaires (SSR et AFLP) ont montré un regroupement de *C. paradisi* avec *C. maxima* (Barkley *et al.*, 2006; Pang *et al.*, 2007). Les données chloroplastiques (Li *et al.*, 2010), ont montré que *C. maxima* est le parent maternel de *C. paradisi*. Ce résultat a été récemment confirmé par des données de séquençage de génomes complets chloroplastiques (Carbonell-Caballero *et al.*, 2015). D'autre part, les données issues de la méthode SCAR (Nicolosi *et al.*, 2000), RFLP, RAPD (Federici *et al.*, 1998) et SSR (Pang *et al.*, 2003), ont montrées un regroupement entre *C. paradisi* et *C. sinensis* et les données RFLP, ont suggéré l'origine paternelle *C. sinensis* de *C. paradisi* (Li *et al.*, 2010). Nos analyses de l'espaceur *trnL-trnF*, montrent *C. paradisi* regroupé avec *C. aurantium*, *C. sinensis*, *C. limon*, *C. tangelo* et *C. maxima* ce qui explique la contribution directe et/ou indirecte de ces espèces comme ancêtres de *C. paradisi*.

C. tangelo (la tangelo 'Minneola') est issue d'une recombinaison entre le patrimoine génétique de *C. reticulata* ('Dancy') et *C. paradisi* (Webber *et al.*, 1943). Nos données de séquences chloroplastiques ont montré *C. tangelo* ('Minneola') regroupée avec *C. aurantium*, *C. sinensis*, *C. limon*, *C. paradisi* et *C. maxima*. Ces résultats suggèrent que l'une de ces espèces est le parent maternel de *C. tangelo* (Figure 32).

2.2.4 Analyse du pseudogène *trnF*_(GAA)

Dans la présente étude l'espaceur intergénique *trnL-trnF* des espèces appartenant au genre *Citrus* montre la présence d'une copie du pseudogène originaire du gène *trnF*_(GAA) chez les 27 variétés du genre *Citrus* au niveau de la position 275pb. En effet, chez différentes espèce du *Solanum*, deux à quatre copies de ce pseudogène ont été identifiées (Poczai & Hyvönen, 2011) (Figure 33). La taille du pseudogène identifié dans la présente étude varie de 49 à 63 pb. Le domaine anticodon est l'élément le plus conservé du pseudogène, quant à la queue acceptrice du côté 3'- est totalement absente. Le domaine D est presque intact; nous avons observé une seule transition (T→C) au niveau de la position 289 pb chez toutes les espèces du genre *Citrus*. Le domaine T paraît très polymorphe entre nos espèces. Cependant, une mutation transversionnelle (T→A) a été identifiée chez les trois cultivars de l'espèce *C. paradisi* à la position 334pb et une transversion (T→G) a été identifiée chez les pomelos 'Ruby' et 'Marsh' observées respectivement à la position 337pb. En outre, nous avons identifié au niveau du même domaine chez tous les cultivars de l'espèce *C. sinensis* et le pomelo 'Star Ruby' de l'espèce *C. paradisi*, une insertion de trois

bases (AAT). Une autre insertion avec l'addition de quatre bases (AAAT) a été également observée au niveau du domaine partiel T de la séquence de 'Marisol' appartenant à l'espèce *C. clementina*. Des délétions ont été aussi observées au niveau du domaine T, allant de 4 ('Nova' et 'Encore' variétés de *C. reticulata* et 'Hernandina' de *C. clementina*) à 15 ('Ortanique' de *C. reticulata*) paires de base à la fin de chaque séquence, de même, la queue acceptrice du côté 5'- a montré deux délétions en amont des séquences.

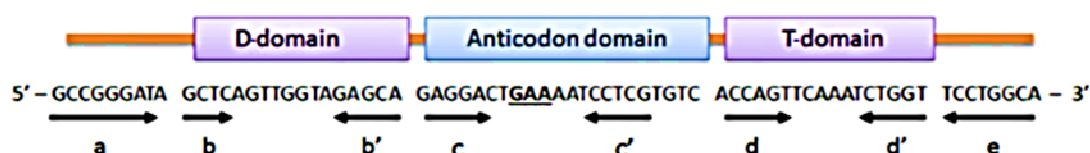


Figure 33 : La séquence nucléotidique du gène *trnF*_(GAA) (Pczai & Hyvönen, 2011)

Les cadres représentatifs forment les éléments structuraux, les flèches du côté 5'- et 3'- représentent le domaine accepteur, les symboles en caractère gras **a-d** et **a'-d'** montrent la structure secondaire

Malgré les délétions observées au niveau de la région acceptrice 5'- elle demeure néanmoins bornée par des motifs observés chez toutes nos séquences. Parmi ces motifs nous avons identifié une importante similarité avec des motifs promoteurs bactérien de type sigma70- (-35 TTGACA; Koch *et al.*, 2007) et de type -10 GAGAT/ -35 TGGGA. En effet, nous avons enregistré deux transitions (A<-->T) en aval de la séquence promotrice. Ces motifs promoteurs ont été séparément identifiés chez l'intergène *trnL-trnF* respectivement aux positions 323 et 91. Poczai & Hyvönen (2011) ont observé ces motifs sans substitution ainsi que d'autres éléments (CTGGGT/CTGGGTTG) bornant le bout du côté 3'- de la copie intermédiaire. Une étude antérieure est arrivée à conclure que le motif -35 TTGACA représente le promoteur original du gène *trnF*_(GAA) (Quandt *et al.*, 2004), toutefois ce promoteur est non fonctionnel car selon Kanno & Harai (1993) il est co-transcrit avec le gène *trnL*_(UAA) (Figure 34).

Il est important de rappeler que les Pseudogènes fournissent un support pour les changements génomiques au niveau de l'ADN. En effet, leur présence est un indicateur de pression évolutive. Ainsi, la formation de pseudogène permet d'accumuler des mutations géniques aléatoires durant l'évolution (Lee, 2003). En outre, le nombre de pseudogène observé indique le taux d'évolution, en effet, un nombre élevé de pseudogènes favorise un haut taux d'évolution. Dans le travail ci présent, nous avons enregistré une seule répétition du pseudogène relativement conservée au sein de nos espèces. L'espaceur intergénique *trnL-trnF* a subi une pression évolutive durant l'histoire des *Citrus* qui pourrait être avec un faible taux. Une étude de l'évolution moléculaire est par conséquent nécessaire afin de juger l'évènement évolutif.

	Séquence du gène <i>trnF</i> _(GAA)	<Acc.Stem> <----D-domain----> <-----Anticodon-----> <----T-domain----> <Acc. Stem>
		GCCGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATCTGGT--TTCCTGGCA
<i>C.aurantium</i>	Bigaradier du Maroc	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAAT-----
	Bigaradier	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAA-----
<i>C.limon</i>	Feminello	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAA-----
	Monachello	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAA-----
	Santa Téréza	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAA-----
	Interdonato	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAA-----
<i>C.paradisi</i>	Pomelo Ruby	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAGG-----
	Pomelo Marsh	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAGG-----
	Pomelo Star Ruby	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAAT-----
<i>C.sinensis</i>	Tarocco	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAAT-----
	Valencia Late	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAATA-----
	Maltaise Barlerin	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAATA-----
	Maltaise 1/2 sanquine	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAATA-----
	Navel Late	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAATA-----
	Lane Late	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAATA-----
	New Hall	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAATA-----
	Navelina	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAATA-----
	Washington Navel	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAATA-----
<i>C.clementina</i>	Cassar	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAATA-----
	Marisol	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATAATA-----
	Hernandina	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATA-----
	MA3	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCA-----
<i>C.tangelo</i>	Minneola	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGT-----
<i>C.reticulata</i>	Ortanique	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACC-----
	Nova	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAATA-----
	Fortune	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAAT-----
	Encore	--CGGGATA--GCTCAGTTGGTAGAGCA--GAGGACTGAAAATCCTCGTGTC--ACCAGTTCAAAT-----

Figure 34 : Alignement manuel des pseudogènes des espèces tunisiennes du genre *Citrus* démontrant leur structure et le polymorphisme observé entre les différentes espèces

2.3 Evolution moléculaire

2.3.1 Les tests de Tajima's et de Fu & Li

Afin d'élucider l'histoire évolutive de la région non codante de l'ADN chloroplastique, deux tests de neutralité ont été appliqués : le test D de Tajima (1989) et le test de Fu & Li (1997). Ces tests permettent d'évaluer l'hypothèse nulle et de détecter la signature de la sélection. Dans le présent travail, nous avons identifié une valeur du test D de Tajima positive et non significatif ($D = 0,48651$, $P > 0,10$) (Tableau 8). De même, des quantités positives et non significatives de D^* et F^* de Fu & Li (1993) ont été obtenues ($D^* = 0,82034$, $F^* = 0,83845$, avec $P > 0,10$ pour les deux tests) (Tableau 8). Ces tests sont basés sur les fréquences alléliques qui sont représentées sur un spectre des fréquences des sites au niveau de l'espaceur intergénique *trnL-trnF* (Figure 35) et sont caractérisés par leur sensibilité aux singletons par rapport au test D de Tajima.

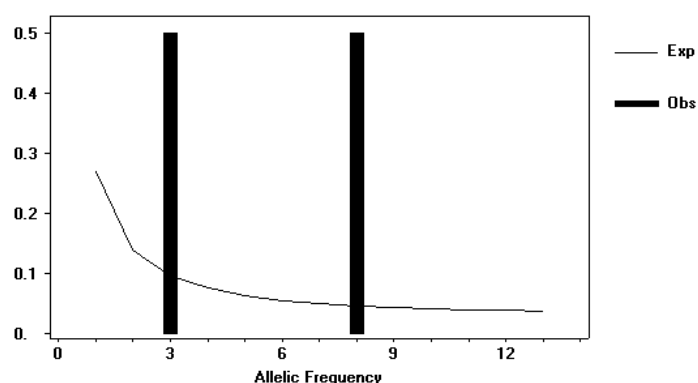


Figure 35 : Spectre des fréquences des sites au niveau des séquences chloroplastiques de l'espaceur intergénique *trnL-trnF* chez les *Citrus* tunisiens

Ainsi, les tests de Tajima et de Fu & Li prouvent un déficit des singletons et révèlent une histoire stable des *Citrus* tunisiens. Ces résultats plaident en faveur d'un scénario démographique stable et confirment l'hypothèse de la neutralité sélective de toutes les mutations (Kimura, 1983).

Tableau 8 : Tests de neutralité sélective au niveau des séquences chloroplastiques de l'espaceur intergénique *trnL-trnF* chez les espèces tunisiennes du genre *Citrus*

Test	Espaceur intergénique <i>trnL-trnF</i>
D de Tajima	0,48651 NS ($P > 0,10$)
D^* de Fu & Li	0,82034 NS ($P > 0,10$)
F^* de Fu & Li	0,83845 NS ($P > 0,10$)
Fs statistique de Fu	0,485 $P = 0,297$

2.3.2 La statistique de Fu

La statistique de Fu suggère que le test Fs est un indicateur plus précis de l'expansion démographique et du hitchhiking que celui de Tajima. Cependant, nos résultats

montrent une valeur de F_s positive et faible ($F_s = 0,485$; $P = 0,678$) et indique une déviation non significative par rapport à la neutralité (Tableau 8). Ces résultats sont significatifs de déficit en allèles entraînant ainsi une perte considérable de la diversité génétique avec une extinction de nombreux variants génétiques. Ces résultats suggèrent la présence d'un cas de goulot d'étranglement et rejettent, par conséquent, le hitchhiking en supportant la neutralité dans l'espaceur intergénique *trnL-trnF*. La distribution de la courbe des Mismatches des séquences révèle une population de taille constante (Figure 36). Les valeurs obtenues basées sur le calcul de l'indice d'Harpending r et de la R^2 statistique ($r = 0,0936$ et $R^2 = 0,1595$) ne suggèrent pas un modèle d'expansion démographique et supportent l'hypothèse de population de taille constante. En effet, les valeurs non significatives des paramètres des tests de neutralité ne confirment pas l'expansion démographique des espèces tunisiennes du genre *Citrus* et supportent bien l'hypothèse de la taille constante de l'échantillon.

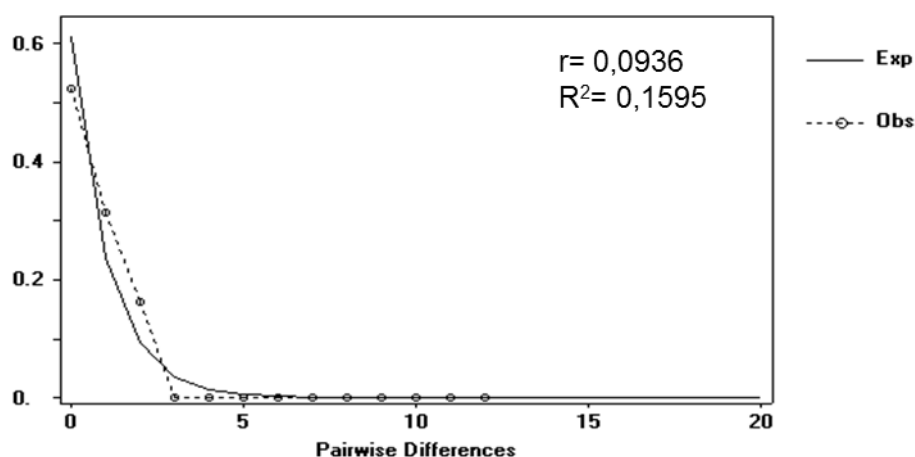


Figure 36: Distribution des Mismatches des séquences chloroplastiques de l'espaceur intergénique *trnL-trnF* chez les *Citrus* tunisiens

2.3.3 Analyse de la différenciation génétique

Afin de mieux comprendre ce qui caractérise et contribue à la répartition observée des espèces tunisiennes du genre *Citrus* sur l'arbre NJ, nous avons étudié la différenciation génétique des agrumes. Pour ce faire, en se basant sur la topologie du dendrogramme NJ (Figure 31), nous avons divisé les agrumes en deux groupes : le premier comprend les variétés des espèces *C. reticulata* et *C. clementina* et le second groupe est composé de tout le reste des espèces (*C. limon*, *C. aurantium*, *C. tangelo*, *C. paradisi* et *C. sinensis*). En outre, cette étude s'est basée sur différents paramètres : le paramètre G_{ST} basé sur la fréquence des haplotypes, le paramètre N_{ST} spécifique des relations génétiques au sein des haplotypes et le coefficient de différenciation entre groupes de Wright appelé F_{ST} . Afin d'étudier la structure de la diversité génétique des espèces tunisiennes du genre *Citrus*,

nous avons adopté la méthode proposée par Pons & Petit (1996) qui consiste à comparer la différenciation entre groupes avec la différenciation génétique en prenant en compte la similarité entre les haplotypes. D'où la comparaison entre le F_{ST} et le N_{ST} . Ainsi, nos applications dans la présente étude montrent une valeur du N_{ST} (0,80542) supérieure à celle du F_{ST} (0,80519) présumant, ainsi, la présence d'un signal phylo-géographique et d'une diversité génétique structurée géographiquement ; les haplotypes phylo-génétiquement les plus proches se retrouvent dans les mêmes groupes. Nous avons adopté une autre approche proposée par Petit *et al.* (2005) qui consiste à comparer entre le N_{ST} et le G_{ST} , paramètres basés respectivement sur la similarité et la fréquence des haplotypes. Ainsi, nous avons enregistré une valeur de G_{ST} (0,65880) inférieure à celle du N_{ST} (0,80542) suggérant la présence d'une structure phylo-géographique et confirmant la première approche.

La valeur du F_{ST} (0,80519) indique une différenciation génétique élevée au sein des deux groupes prédéfinie au niveau de l'espaceur intergénique *trnL-trnF* (Wright, 1978). La valeur du G_{ST} supérieure à 0,5 suggère une diversité élevée expliquée par un faible flux génique entre les deux groupes. La valeur faible du paramètre N_m (0,13 ; estimé en se basant sur le G_{ST}) supporte cette hypothèse.

3. Conclusion

Dans la présente étude, les séquences de l'espaceur intergénique *trnL-trnF* montrent un polymorphisme moléculaire considérable entre les différentes espèces tunisiennes du genre *Citrus* avec de faible flux génique entre les *Citrus*. Nos résultats ont permis d'élucider les liens de parentés entre les différentes espèces du genre *Citrus*. En effet, nous avons enregistré un regroupement important entre les espèces *C. aurantium*, *C. sinensis*, *C. limon*, *C. tangelo*, *C. paradisi* et *C. maxima*. Ce rassemblement suggère que *C. maxima* est considéré comme l'ancêtre commun des autres espèces du même groupe. D'autre part, les données chloroplastiques révèlent l'origine maternelle de *C. clementina* qui est *C. reticulata*. En plus, nos résultats ont prouvé la contribution directe et indirecte de l'espèce *C. maxima* comme parent maternel du groupe des pamplemousses. Nos résultats montrent clairement la contribution des pools génétiques du pamplemousse et de la mandarine dans le développement des espèces secondaires tunisiennes. D'autre part, le pseudogène *trnF*_(GAA) détecté au niveau des séquences chloroplastiques est le résultat d'accumulation de mutations au cours de l'évolution. En revanche, les valeurs des tests de la neutralité sélective s'écartent non significativement du zéro, la population est donc soumise à un goulot d'étranglement avec un faible taux de mutation constant au cours du temps et un faible niveau de la diversité nucléotidique. Par conséquent, nous suggérons que ce locus a évolué sous un modèle de mutations strictement neutre avec la présence d'un équilibre entre

la mutation et la dérive génétique (Simonsen *et al.*, 1995). A vrai dire, le nombre des allèles perdus par la dérive est égal au nombre d'allèles produits par la mutation. En outre, l'obtention d'une population de taille constante supporte cette hypothèse. Une population qui a subi un goulot d'étranglement est une population qui a perdu des haplotypes, ce qui explique le faible nombre d'haplotypes enregistré au niveau des séquences de l'espaceur intergénique *trnL-trnF*. Il est important de noter que la théorie neutraliste prédit que le polymorphisme et la divergence ne dépendent que de la taille effective de la population et du taux de mutation. Le taux de recombinaison n'affecte pas la neutralité et la divergence entre les espèces, mais il est en corrélation avec le polymorphisme. En effet, l'absence d'évènement de recombinaison montre un faible niveau de polymorphisme.

Finalement, en se basant sur l'espaceur intergénique *trnL-trnF*, la biodiversité des variétés tunisiennes montre une importante organisation au niveau moléculaire et indique que les échanges génétiques entre les taxons sont limités. L'apomixie partielle associée à la polyembryonie de la plupart des espèces secondaires a été certainement un élément essentiel dans la limitation du flux de gènes entre les *Citrus* considérés (Ollitrault & Navarro, 2012). La présente étude a démontré que les données des séquences *trnL-trnF* sont d'une puissante résolution dans la démonstration de la contribution des pools génétiques du pamplemousse et de la mandarine, dans le développement des espèces secondaires tunisiennes du genre *Citrus* et dans la déduction de leur lien génétique. Nous avons tracé l'origine chloroplastique des espèces tunisiennes du genre *Citrus* et des études récentes basées sur les marqueurs moléculaires ciblant des codons d'initiation (SCoT : *Start Codon Targetted*) ont tracé l'origine nucléaire (Mahjbi *et al.*, 2015).

Chapitre II : Elaboration d'une clé moléculaire taxonomique de la sous famille des Aurantioideae en utilisant les marqueurs moléculaires SNPs diagnostiques des principaux clades génotypés par la PCR compétitive allèle-spécifique

1. Identification des SNPs des clades spécifiques

L'identification des SNPs des clades spécifiques a été élaborée en suivant une procédure bien précise. Ainsi, les 79 séquences des variétés sélectionnées de la sous famille des Aurantioideae de huit régions chloroplastiques ont été alignées et organisées selon la classification de Bayer *et al.* (2009) (Figure 38 ; Annexe 1). Ces séquences ont été alignées à la séquence de référence du génome chloroplastique complet de *C. sinensis* (Bausher *et al.*, 2006) dans le but de localiser les SNPs sélectionnés dans ce génome de référence. Ce qui a permis une simple identification visuelle des SNPs différenciant totalement les différents clades du reste des variétés.

Un exemple d'une telle identification est montré au niveau de l'Annexe 4 montrant une identification d'un SNP diagnostique du clade Q (celui de Bayer *et al.*, 2009) localisé au niveau de la région de l'espaceur *rps4-trnT* et qui correspond au groupe des agrumes vrais de Swingle & Reece (1967) et les genres *Oxanthera* et *Feroniella*. En effet, au niveau de la position 49 301 de la séquence de référence du génome chloroplastique, un SNP (A/G) différencie totalement les variétés du clade Q (A) des autres variétés de la sous famille des Aurantioideae (G).

Par ailleurs, sur les 54 clades analysés un total de 166 SNPs diagnostiques a été identifié. Parmi lesquels huit n'ont révélé aucun SNP diagnostique (dont un affiche un SNP partagé avec le clade GG1 et les espèces : *C. aurantifolia* et *Hesperthusa crenulata*), la majorité représentée par 28 clades révèle entre 1 et 3 SNPs diagnostiques, 15 clades révèlent entre 4 et 8 SNPs diagnostiques et seulement 3 révèlent 12 ou plus (Figure 37 et Tableau 9).

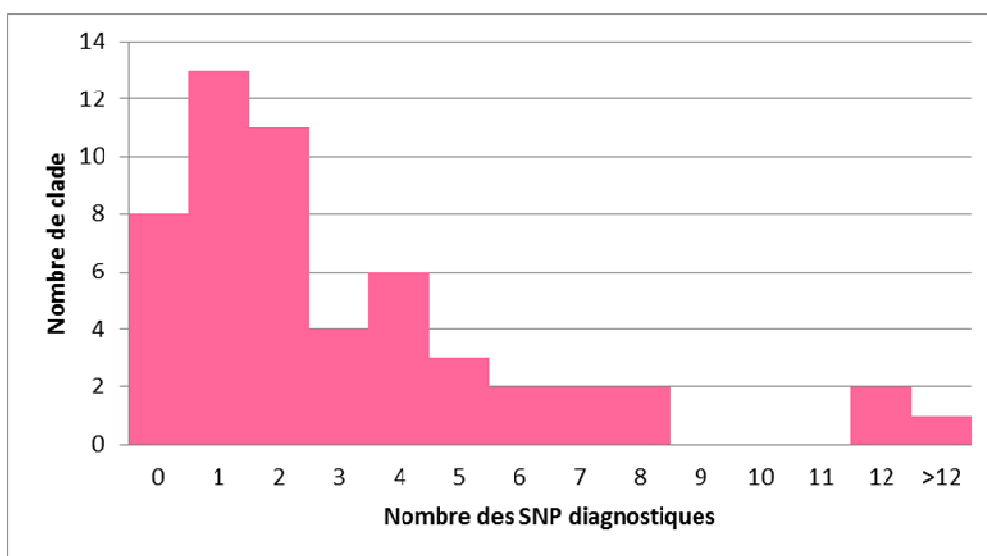


Figure 37 : Distribution du nombre des SNPs diagnostiques au niveau des 54 clades considérés de Bayer *et al.* (2009)

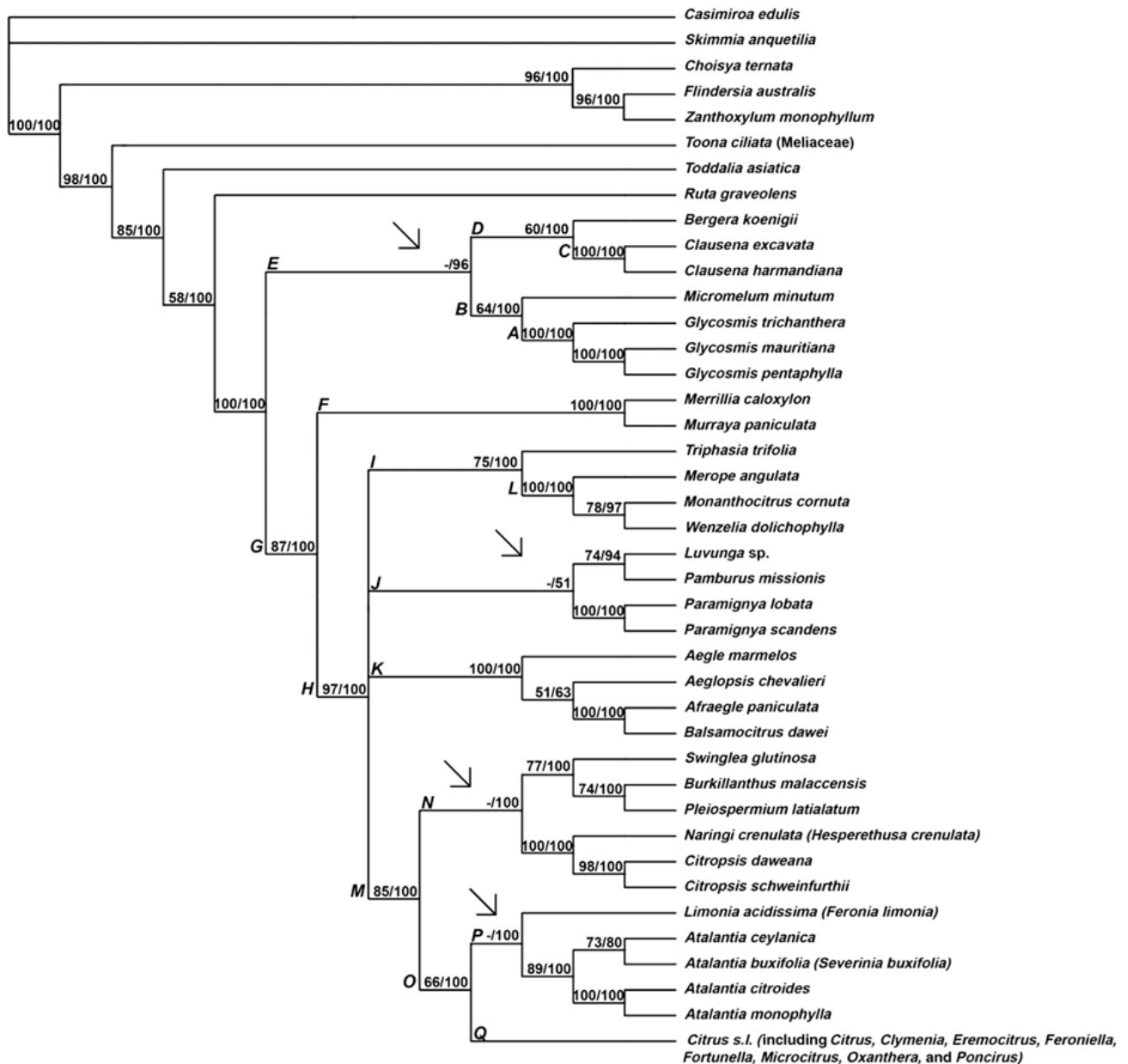
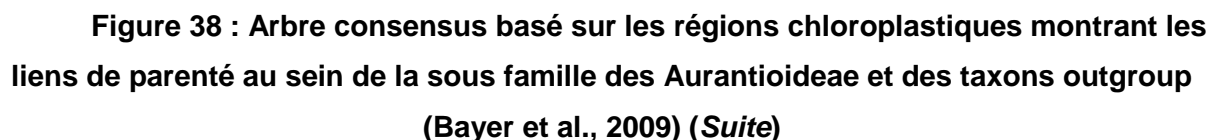


Figure 38 : Arbre consensus basé sur les régions chloroplastiques montrant les liens de parenté au sein de la sous famille des Aurantioideae et des taxons outgroup (Bayer *et al.*, 2009)



Ainsi, les valeurs élevées de 12 SNPs diagnostiques ou plus par clade ont été identifiées sur *Poncirus trifoliata* et *Clymenia polyandra*, espèces représentatives respectivement des clades RR et PP. Ces valeurs sont partiellement expliquées par le fait que ces taxons ont été représentés par un seul échantillon entraînant ainsi une perte de la diversité intra-taxon. Néanmoins, ceci témoigne d'une importante différenciation de ces genres au sein du clade des agrumes vrais. Alors que le clade A, troisième clade représentatif des valeurs élevées des SNPs diagnostiques, est représenté par trois espèces du genre *Glycosmis*. Ce qui confirme la conclusion de Bayer *et al.* (2009) et Samuel *et al.* (2001) sur la robustesse du monophylétisme des *Glycosmis*.

L'identification de 8 SNPs diagnostiques pour le clade F (*Merilla* et *Murraya paniculata*) et K (*Aegle*, *Aeglopsis*, *Afraegle* et *Balsamocitrus*) et 7 pour le clade Q (les agrumes vrais avec *Feroniella* et *Oxanthera*) valide la robustesse du caractère monophylétique de chaque clade. En outre, pour le clade F, ces résultats confirment le lien étroit de *Merilla* et *Murraya paniculata* proposé par Bayer *et al.* (2009). Le clade K a été défini par Bayer *et al.* (2009) comme le clade des Balsamocitrinae qui correspond au groupe des fruits Baels de Swingle & Reece (1967). De même, l'origine monophylétique du clade Q, qui inclut les séquences des agrumes vrais (Swingle & Reece, 1967), *Feroniella* et *Oxanthera*, est aussi clairement identifiée par les valeurs des SNPs diagnostiques. De plus, avec 7 SNPs diagnostiques *C. macroptera* apparaît clairement différenciée des autres papeda comme *C. micrantha* et *C. celebica* pourtant elles ont été classées dans le clade DD validé par 2 SNPs diagnostiques. Quant aux clades OO de l'espèce *Oxanthera* et V des espèces *Microcitrus australasica* et *Microcitrus garraway* sont validés par 6 SNPs diagnostiques ce qui révèle une importante différenciation au sein du clade Australien/Néo-Calédonien (clade T). Avec 5 SNPs diagnostiques le clade QQ des espèces *C. medica* et *C. indica* est clairement identifié. Mais le clade R, qui regroupe les clades QQ et T (Australien/Néo-Calédonien), a été identifié par un seul SNP diagnostique. Ce résultat révèle un lien étendu entre ces deux groupes, bien que ce clade (R) a été aussi déduit des données des séquences des génomes chloroplastiques complets (Carbonnel-Caballero *et al.*, 2015). L'identification de 5 SNPs diagnostiques pour les clades E et G supporte l'origine monophylétique de *Bergera/Clausena/Micromelum/Glycosmis* d'un côté et de tous les autres genres d'un autre côté. Aussi avec 5 SNPs diagnostiques *C. ichangensis* apparaît clairement identifiée de toutes les papedas. Avec 4 SNPs diagnostiques, les origines monophylétiques des clades C (*Clausena*), D (*Bergera* et *Clausena*), L (*Merope/Monanthocitrus/Wenzelia*) et U (*Citrus* Australien) ont été aussi clairement validées.

L'absence de SNP diagnostique chez plusieurs clades correspond à la faible robustesse des nœuds de l'arbre consensus phylogénétique préparé par Bayer *et al.* (2009) et souligne la faiblesse de ces clades. Ceci correspond à l'origine monophylétique de

différents clades : citant en premier lieu 'i', le clade I regroupant *Triphasia trifolia* et le clade *Merope/Monanthocitrus/Wenzelia*, en second lieu 'ii', le clade O: (*Feronia* et *Atalantia*) et (agrumes vrais avec *Oxanthera* et *Feroniella*).

Tableau 9 : Nombre des SNPs diagnostiques des clades de Bayer et al. (2009)

	région codante <i>atpB</i>		espaceur <i>rbcL-atpB</i>		espaceur <i>rps16</i>		région <i>trnL-F</i>		espaceur <i>rps4-trnT</i>		espaceur <i>matK-5'trnK</i>		espaceur <i>psbM-trnD^{GUC}</i>		intron <i>trnG</i>		Total	
Clade Bayer	D	D'	D	D'	D	D'	D	D'	D	D'	D	D'	D	D'	D	D'	D	D'
A	0	0	0	0	1	0	1	2	0	0	4	2	3	1	3	0	12	5
B	0	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	3	0
C	0	0	1	0	1	0	0	2	0	0	0	0	0	0	2	0	4	2
D	1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0
E/G	0	0	1	0	0	0	0	1	3	0	1	0	0	0	0	0	5	1
F	2	0	1	1	1	0	0	0	1	0	0	0	3	0	0	1	8	2
H	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	2	0
I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	1
K	2	0	0	0	2	3	0	0	0	0	1	0	0	0	3	0	8	3
L	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0
L1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L2	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	4	0
L3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
M	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	2	1
N	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0	2	1
Q	0	2	1	1	0	0	0	0	1	0	1	0	0	0	4	0	7	3
R	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
S	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U	0	0	0	0	1	0	0	0	0	0	1	0	1	0	1	0	4	0
V	0	0	0	0	0	0	0	0	1	0	1	0	2	0	2	1	6	1
W	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0
X	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	2	1
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Z	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	3	0
AA	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	2	0
BB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
CC	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	1	3
DD	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	2	1
EE	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	3	0
FF1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0
FF2	0	0	0	0	0	0	3	0	1	0	1	0	1	0	1	1	7	1
GG1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
GG2	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0

Table 9: Nombre des SNPs diagnostiques des clades de Bayer *et al.* (2009) (Suite)

	région codante <i>atpB</i>		espaceur <i>rbcL-atpB</i>		espaceur <i>rps16</i>		région <i>trnL-F</i>		espaceur <i>rps4-trnT</i>		espaceur <i>matK-5'trnK</i>		espaceur <i>psbM-trnD^{GUC}</i>		intron <i>trnG</i>		Total	
	D	D'	D	D'	D	D'	D	D'	D	D'	D	D'	D	D'	D	D'	D	D'
Clade Bayer																		
GG3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	2	0
GG4	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0
GG5	1	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	2
GG6	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2	0
HH	1	0	1	0	0	0	0	0	0	0	0	0	2	0	1	0	5	0
II	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	2	2
JJ	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	1
KK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	0
LL	0	0	0	0	1	0	0	0	1	0	0	0	1	0	1	0	4	0
MM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
OO	1	0	1	1	0	1	1	0	1	1	1	2	1	0	0	0	6	5
PP	3	0	2	3	2	0	2	0	0	0	1	0	3	0	1	0	14	3
QQ	0	0	2	0	2	1	0	0	0	0	1	0	0	0	0	0	5	1
RR	1	1	1	0	1	1	2	0	0	0	3	0	2	1	2	1	12	4
SS	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0
TT	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Total	19	5	20	7	19	8	15	6	13	1	20	9	28	4	32	6	166	46

Avec : D: marqueur totalement diagnostique du clade, D': marqueur partagé par toutes les variétés du clade et quelques variétés en dehors du clade

Il semble que la résolution des sous-groupes de la tribu des Citreae demeure confuse et les données des séquences des génomes chloroplastiques complets seront probablement nécessaires pour conclure définitivement sur l'organisation de cette tribu. Ainsi, l'origine monophylétique des sous-groupes a été aussi mise en question : 'iii' celle du clade T: *Clymenia polyandra* et le clade Australien/Néo-Calédonien, 'iv' et celle du clade Y regroupant les clades BB (mandarines) et Z (pamplemousses et papedas). Les autres branchements identifiés sans SNP diagnostique concernent des différenciations récentes au sein de variétés très proches (L1; GG1; MM; NN).

Finalement, 40 SNPs diagnostiques de 33 clades parmi les 166 SNPs diagnostiques identifiés ont été sélectionnés. Pour chaque clade, le principale critère de sélection repose sur l'absence de polymorphisme additionnel (SNP ou InDel) dans la proximité du SNP. Pour ce faire, une absence de polymorphisme additionnel au niveau des 30pb flquant le SNP de chaque côté serait idéale, mais au cas où pour certains clades cette condition n'est pas réalisable, se limiter à 15 pb serait aussi efficace.

2. Analyse KASPar

Les amorces des 40 SNPs sélectionnés ont été construites par le LGC Genomics® en se basant sur les séquences flanquantes du locus SNP de, approximativement, 50 nucléotides de part et d'autre du SNP (Annexe 3). Parmi ces amorces, 11 étaient inutiles ; certains n'ont fourni aucun signal d'autres ont donné des signaux incohérents : CP_3707 (clade DD), CP_10192 (clade U), CP_32282 (clade N), CP_5927 (clade EE), CP_51032 (clade II), CP_32684 (clade Z), CP_57022 (clade A), CP_58121 (clade J), CP_51540 (clade JJ), CP_57256 (clade L) et CP_3996 (clade Q). Pour le marqueur CP_57686, théoriquement diagnostique de « *Ichang papeda* », a fourni un bon signal mais non polymorphe, il a été par conséquent supprimé. En outre, le marqueur CP_57742 théoriquement diagnostique de *Clymenia* différencie *Atalantia ceilanica*, et non *Clymenia*, de toutes les autres variétés a aussi été par conséquent retiré. Par la suite, les analyses génétiques ont été effectuées avec les 27 marqueurs diagnostiques restants des 24 clades de Bayer *et al.* (2009) testés sur 108 variétés (Tableau 10). Un exemple de l'application des marqueurs KASPar décrit au niveau par la Figure 39 illustre une différenciation nette obtenue avec le marqueur CP_49301 (A/G) entre les variétés des agrumes vrais (selon la nomenclature de Swingle & Reece, 1967), *Oxanthera* et *Clymenia* (variétés représentatives du clade Q) avec tous les autres genres.

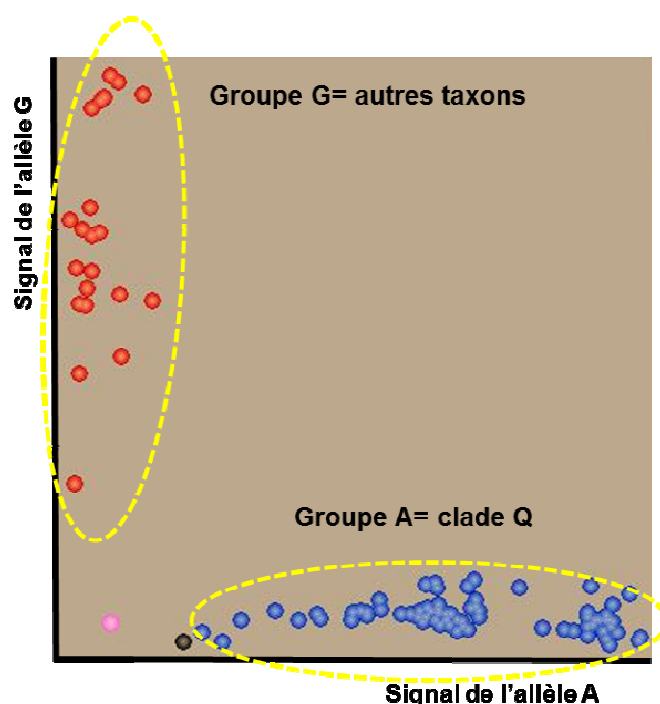


Figure 39 : Un exemple de l'analyse KASPar du marqueur CP_49301 (A/G) distinguant les variétés du clade Q de Bayer *et al.* (2009)

Avec: ● Allèle G, ● Allèle A, ● Témoin négatif, ● Donnée manquante

Le pourcentage des données manquantes de ces marqueurs est très faible, il varie de 0 à 2,8% avec une moyenne de 1,2%. Cette situation est très différente de celle affrontée avec les données des marqueurs SNPs KASPar nucléaires extraites d'un ensemble de variétés d'agrumes où la fréquence des données manquantes était élevée pour les genres proches des agrumes et augmentait avec la distance taxonomique au sein de la sous famille des Aurantioideae (Garcia-Lor *et al.*, 2013). Ce qui prouve que la sélection basée sur l'absence de polymorphisme additionnel aux alentours du SNP ciblé était efficace. Cette très bonne transférabilité résulte aussi du faible niveau évolutif des séquences chloroplastiques par rapport aux séquences nucléaires. D'autre part, le PIC varie de 0,02 à 0,5 en relation avec le niveau taxonomique révélé par chaque marqueur mais aussi dû à notre échantillon surreprésenté par les agrumes vrais par rapport aux autres genres.

Tableau 10 : Résultats des analyses des marqueurs KASPar

Marqueur	Position	Ref/Alt	Clade	% MD	CDi	PIC
CP_58345	58345	G/A	QQ	2,8%	1,00	0,11
CP_4296	4296	G/A	QQ	0,9%	1,00	0,11
CP_32299	32294	G/T	R	4,6%	1,00	0,28
CP_58401	58401	G/T	S	1,8%	1,00	0,50
CP_56697	56697	T/C	D	0,0%	1,00	0,07
CP_3807	3807	C/T	E	1,8%	1,00	0,09
CP_58454	58454	C/T	G	1,8%	1,00	0,09
CP_32526	32521	A/G	H	1,8%	1,00	0,12
CP_3964	3964	C/A	Eremocitrus	1,8%	1,00	0,02
CP_9149	9149	T/A	LL	0,9%	1,00	0,02
CP_58062	58062	T/G	X	0,0%	0,98	0,15
CP_10342	10343	C/A	AA	0,9%	0,96	0,15
CP_51634	51634	C/A	FF	0,0%	1,00	0,09
CP_3674	3674	C/G	V	0,0%	1,00	0,05
CP_4243	4243	A/C	W	0,0%	1,00	0,10
CP_57691	57691	T/C	RR	0,0%	1,00	0,04
CP_4159	4159	A/C	CC	0,9%	0,92	0,36
CP_57064	57064	G/A	F	2,8%	1,00	0,04
CP_32463	32458	T/C	J	0,0%	1,00	0,04
CP_9158	9158	T/C	K	0,9%	1,00	0,09
CP_3829	3829	T/C	M	1,8%	0,88	0,23
CP_49405	49405	A/G	M	2,8%	0,98	0,24
CP_3747	3747	G/A	N	0,0%	0,73	0,12
CP_51068	51068	A/C	OO	0,0%	1,00	0,02
CP_3982	3982	C/T	P	1,8%	0,96	0,12
CP_49301	49301	A/G	Q	0,9%	1,00	0,41
CP_58553	58553	G/T	Q	0,0%	1,00	0,41

Clade: Clade différencié par le SNP accordé aux analyses *in silico*, %MD: pourcentage des données manquantes, Cdi: *Clade differentiation rate* (G_{ST}), PIC: *Polymorphism Information Content*

L'analyse du G_{ST} a été effectuée en considérant le clade théoriquement diagnostiqué par le SNP comme une population et le reste des échantillons comme une autre population. Cette analyse a globalement validée la valeur du clade diagnostique des marqueurs sélectionnés avec une valeur moyenne du G_{ST} de 0,98. Ainsi, 20 marqueurs sont totalement fixes avec une valeur du G_{ST} égale à 1.

Le clade CC est représenté par les marqueurs CP_4159 et CP_51032 où 'Yuzu', la mandarine japonaise appartenant à l'espèce *C. junos* (selon la classification de Tanaka), est différente des mandarines alors qu'elle en faisait partie dans les analyses de Bayer *et al.* (2009). Nos résultats sont, néanmoins, en concordance avec les résultats de Abkenar *et al.* (2011) et Yamamoto *et al.* (2013) qui ont trouvé 'Yuzu' différencié des mandarines en ayant recours aux analyses chloroplastiques par les marqueurs moléculaires CAPS. Par conséquent, seulement le statut de 'Yuzu' reste discutable et ces deux marqueurs figurent comme des marqueurs diagnostiques idéaux du groupe des mandarines. Pour le clade M, les deux marqueurs sélectionnés montrent un échantillon en outgroup avec le même allèle que celui du clade M (*Clausena excavata* pour le marqueur CP_3829 et *Severinia disticha* pour le marqueur CP_49405).

Le clade X représenté par le marqueur CP_58062 spécifique de *Fortunella* et *C. halimi* montre aussi une variété (*Atalantia ceylanica*) en dehors du clade partageant le même allèle du clade. Pour le marqueur CP_10342 (clade AA ; *Fortunella*) l'allèle diagnostique est partagé avec deux variétés du clade D (*Clausena anisata* et *Murraya koenigii*) en plus des six espèces *Fortunella* analysées. Pour l'espèce *M. koenigii* cela confirme les résultats obtenus par les analyses des séquences de Bayer *et al.* (2009). De même pour le marqueur CP_3982 spécifique du clade P celui de *Limonia* et *Atalantia*, l'allèle diagnostique a été partagé par deux variétés du clade N (*Hespertusa* et *Pleiospermium*). Pour le marqueur CP_3747 représentatif du clade N, affiche la plus faible valeur du G_{ST} avec une valeur égale à 0,73 où une variété du clade théorique partage un allèle en dehors du clade alors que *Severinia buxifolia* partage l'allèle du clade. Ainsi, les résultats obtenus par les deux derniers marqueurs montrent leur résolution insuffisante du clade des Citrinae dans la cohérence des sous-clades.

L'analyse UPGMA a été effectuée avec 27 marqueurs SNPs (Figure 40). La représentation est hautement cohérente avec les clades obtenus par les analyses de Bayer *et al.* (2009).

Il apparaît que, les six genres du groupe des agrumes vrais, classés selon la nomenclature de Swingle & Reece (1967), sont inclus dans le clade Q de même pour le représentatif du genre *Oxanthera* comme a été proposé par Bayer *et al.* (2009). Cependant, *Feroniella oblata* espèce représentative du genre *Feroniella* a été identifiée au niveau du clade M', étroitement associé aux espèces représentatives de Balsamocitrinae (comme

présenté par la classification de Swingle & Reece (1967)), des groupes des fruits primitifs des agrumes et des fruits proches des agrumes et non pas dans le clade Q comme conclu par Bayer *et al.* (2009). Pour cette raison, la conformité génétique de l'échantillon de *Feroniella* utilisées par Bayer *et al.* (2009) ou dans notre étude pourrait être remise en question.

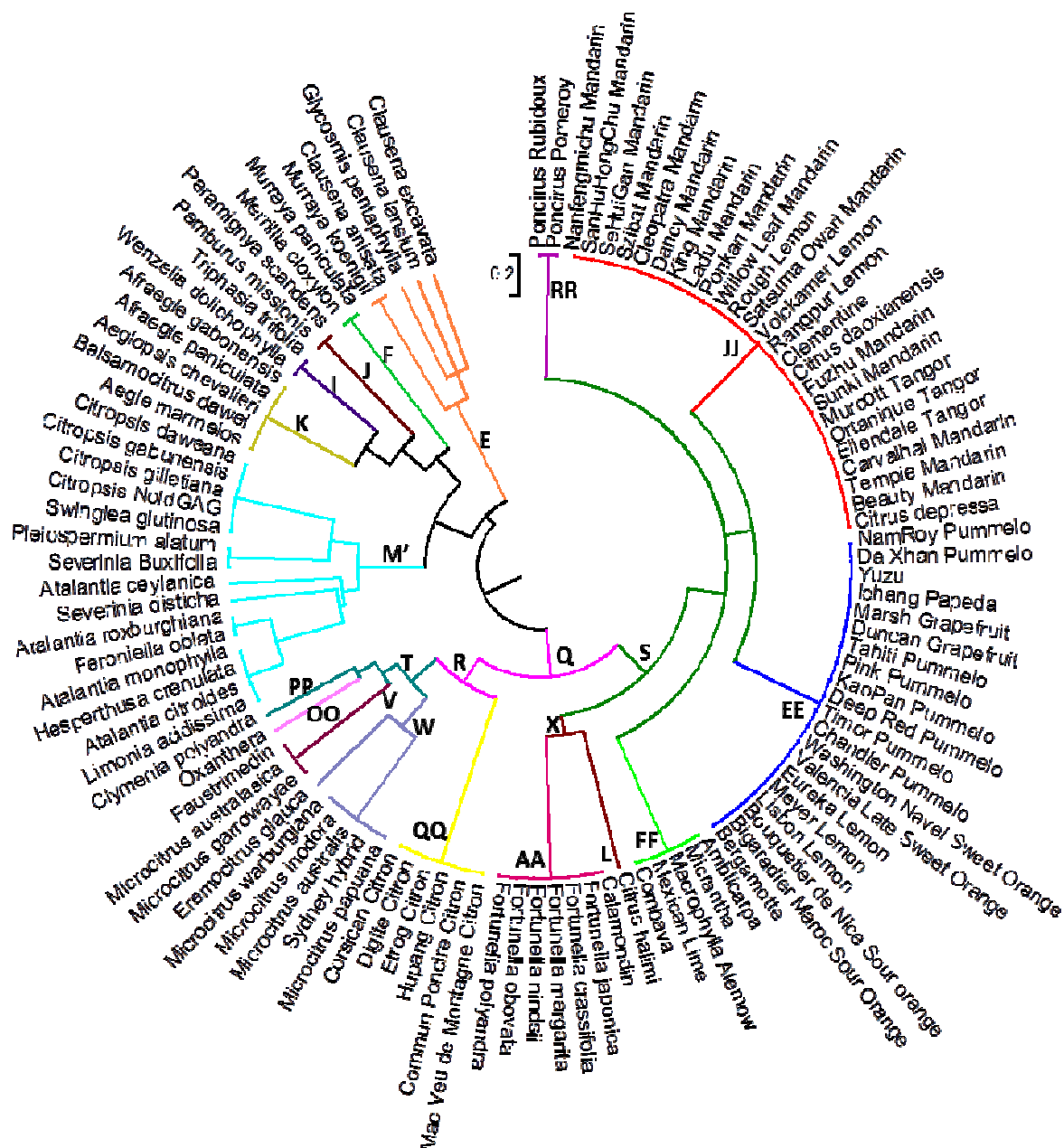


Figure 40 : Classification des variétés de la sous famille des Aurantioideae en utilisant 27 marqueurs SNPs diagnostiques des clades Bayer *et al.* (2009)

Au sein du clade Q, l'organisation générale est globalement conforme avec 15 clades des clades Bayer *et al.* (2009) rencontrés et avec seulement la position de 'Yuzu' et de

'Ichang papeda' dans le clade des pamplemousses étant différente. Le clade RR des *Poncirus* est clairement différencié. Le clade JJ des mandarines est largement étendu à 14 espèces (selon la classification de Tanaka) : *C. depressa*, *C. tangerine*, *C. daoxianensis*, *C. clementina*, *C. kinokuni*, *C. erythrosa*, *C. suhuiensi*, *C. reticulata*, *C. reshni*, *C. nobilis*, *C. paratangerina*, *C. sunki*, *C. deliciosa* et *C. unshiu*. De même, les tangors 'Ortanique' et 'Murcott', supposées être naturelles, partagent ce clade comme a été conclu par Garcia-Lor *et al.* (2015) en aillant recours aux marqueurs SSRs chloroplastiques et aux InDels mitochondriaux. En outre, trois agrumes acides : la lime 'Rangpur', 'Rough lemon' et 'Volkameriana' sont aussi associés à ce clade en concordance avec Bayer *et al.* (2009) pour 'Rangpur' et 'Rough lemon'. Ces trois agrumes acides utilisés comme porte-greffe confirment les conclusions de Froelicher *et al.* (2011) avec les InDels mitochondriaux et les plus récentes dérivées d'une large combinaison des analyses cytoplasmiques et nucléaires démontrant que ces trois limes et citrons résultent d'une hybridation directe entre les pools génétiques de la mandarine et du cédrat (Curk *et al.*, 2016). Avec les données SNPs disponibles, nous ne sommes pas en mesure de développer des marqueurs de différenciation entre mandarine douce et amère observées avec les données mitochondriales (Froelicher *et al.*, 2011), les SSRs chloroplastiques (Garcia-Lor *et al.*, 2015) et les données des séquences des génomes complets (Carbonelle-Caballero *et al.*, 2015). La disponibilité récente des données des séquences des génomes complets va certainement permettre de développer des marqueurs KASPar SNPs pour cette différenciation. Le clade EE des pamplemousses inclus tous les pamplemousses analysés, les oranges amères et douces, les pomelos, les citrons 'Meyer' et 'Lisbon', 'Ichang papeda' et "Yuzu. L'hérédité maternelle pamplemousse de l'orange amère et douce et pomelo a été aussi décrite par plusieurs auteurs (Nicolosi *et al.*, 2000; Asadi Abkenar *et al.*, 2004; Barkley *et al.*, 2006; Tshering Penjor *et al.*, 2010; Li *et al.*, 2010; Froelicher *et al.*, 2011; Yamamoto *et al.*, 2013; Carbonell Caballero *et al.*, 2015). Nicolosi *et al.* (2000) ont proposé *C. limon* comme un hybride entre l'orange amère et le cédrat et a été définitivement démontré par Curk *et al.* (2016) de même pour l'hérédité maternelle pamplemousse de 'Meyer' et 'Bergamot'. Le développement de deux marqueurs diagnostiques SNPs additionnels différenciant (i) entre l'orange amère et le pamplemousse a été déjà révélé par les marqueurs moléculaires chloroplastiques de type SSR (Curk *et al.*, 2016) et les données des séquences des génomes complets (Carbonell Caballero *et al.*, 2015) et (ii) entre l'orange douce, le pamplemousse et l'orange amère a été révélé par les données des séquences des génomes complets (Carbonell Caballero *et al.*, 2015) permettrait une définition très fine de ce clade avec quelques marqueurs SNPs. Le clade FF de la 'lime Mexicaine' inclus aussi la 'Combava' de l'espèce *C. hystrix* de même *C. micrantha*, *C. macrophylla* et *C. amblycarpa*. Ces résultats sont en concordance avec ceux de Froelicher *et al.* (2011), Garcia-Lor *et al.*

(2015) et Curk *et al.* (2016) issus des InDels mitochondriaux et des SSRs chloroplastiques. Les six espèces analysées de *Fortunella* sont identiques (clade AA) et ont été groupées avec *C. halimi* au niveau du clade X conformément aux résultats de Bayer *et al.* (2009). Ces résultats confirment les études de Roose (1998), Scora (1988), les données isozymatiques de Herrero *et al.* (1996), les études chemotaxonomiques de Ogawa *et al.* (2001), les données chloroplastiques issues de la technique RFLP (Gulsen & Roose, 2001) et les données des gènes nucléaires (Ramadugu *et al.*, 2013). Cette association génétique devrait être due à l'origine géographique de *Fortunella* et de *C. halimi* de la Péninsule Malaisienne (Gulsen & Roose, 2001). Tous les cédrats sont regroupés dans le clade QQ qui est lié au clade R avec les agrumes Australiens/Néocalédoniens. Ce lien de parenté entre *C. medica* et les agrumes Australiens/Néocalédoniens a été révélé pour la première fois par Pfeil *et al.* (2008) puis par Bayer *et al.* (2009) et récemment confirmé par les analyses des séquences des génomes complets chloroplastiques (Carbonnel Caballero *et al.*, 2015). Avec le clade T des agrumes Australiens/Néocalédoniens, l'organisation présentée par Bayer *et al.* (2009) est principalement conservée avec les clades V et W.

Le clade M' correspond au clade M de Bayer *et al.* (2009) mais il n'inclut pas le clade Q des agrumes vrais et *Oxanthera*. Ça pourrait être dû à la sélection de deux SNPs diagnostiques pour le clade Q. C'est un clade où quelques SNPs diagnostiques ont été identifiés et sa structuration n'a pas été identique ni à celle présentée par Bayer *et al.* (2009) ni avec celle de Swingle & Reece (1967). Un ajout de plus de marqueurs à ce clade serait nécessaire pour permettre une bonne résolution de ce groupe. De même, Penjor *et al.*, (2013) et Morton *et al.* (2003) n'ont pas supporté la distinction de la sous tribu des Citrinae entre les groupes des fruits des agrumes primitifs et des fruits proches des agrumes classés par Swingle & Reece (1967). En effet, Morton (2009) suggère une révision des tribus et des sous-tribus de la sous famille des Aurantioideae.

Les principaux clades pour le reste des tribus des Citreae et des Clauseneae ont été respectés par rapport à la classification de Bayer *et al.* (2009) avec (i) le clade K regroupant *Afraegle*, *Aeglopsis*, *Balsamocitrus* et *Aegle*; (ii) le clade I associant *Wenzelia* et *Triphasia trifolia*; (iii) le clade J composé de *Pamburus* et *Paramigna*; (iv) le clade F constitué de *Merilla* et *Murraya paniculata* et finalement (v) le clade E des *Clauseneae* formé de *Murra koenigii*, *Clausena* Sps et *Glycosmis penthaphylla*.

3. Utilité et complémentarité de l'analyse KASPar/SNPs diagnostiques par rapport à d'autres marqueurs moléculaires

Bien que l'analyse de l'ADN chloroplastique dans les études phylogénétiques est essentiellement basée sur des régions non codantes aussi bien que l'analyse de données des séquence de gènes (Jung *et al.*, 2005; Li *et al.*, 2007; Jena *et al.*, 2009; Li *et al.*, 2010;

Abkenar *et al.*, 2011), plusieurs études ont analysé la phylogénie maternelle de la sous famille des Aurantioideae en utilisant différents marqueurs moléculaires tel que la RFLP (Olmstead & Palmer, 1994; Green *et al.*, 1986; Yamamoto *et al.*, 1993; Gulsen & Roose, 2001; Abkenar *et al.*, 2004; 2011; Jena *et al.*, 2009), les SSRs chloroplastiques (Deng *et al.*, 2007; Garcia *et al.*, 2015; Hang *et al.*, 2015), les SSCP (Cheng *et al.*, 2005; Olivares-Fuster *et al.*, 2007) et la CAPS (Loffy *et al.*, 2003; Yamamoto *et al.*, 2013). La séquence complète du génome chloroplastique de *C. sinensis* (Bausher *et al.*, 2006) a poussé vers des études approfondies comme l'utilisation de la séquence complète du génome pour élucider les relations phylogénétiques et l'estimation du temps de divergence (Carbonnel Caballero *et al.*, 2015) et a facilité le développement des marqueurs SNPs sur l'ensemble du génome chloroplastique. Les SNPs sont les plus abondants type de variation de séquence chez les génomes eucaryotes et ont été considérés être la plus simple, l'ultime et la plus petite unité de l'hérédité.

Ainsi, les SNPs chloroplastiques ont été identifiés chez l'espèce *Populus* et l'approche CAPS a été développée pour révéler le polymorphisme entre sept espèces (Schroeder & Fladung, 2014). Mais cette approche présente différents inconvénients dont les principaux sont liés au rendement faible et au génotypage relativement cher. D'autre part, un considérable ensemble de SNPs chloroplastiques a été identifié chez la Soja, sélectionné et inclus correctement sur des puces de génotypage (Lee *et al.*, 2015) ainsi que ce fut fait sur la pomme de terre *Solanum tuberosum* (Vos *et al.*, 2015).

Toutefois, l'utilisation des puces à génotypage demeure couteuse et ne permet pas une flexibilité d'adaptation à des questions spécifiques de recherche nécessitant seulement un sous ensemble spécifique limité de marqueurs. Pour la première fois dans la taxonomie végétale, nous avons démontré le concept qu'une bonne sélection d'un ensemble de marqueurs SNPs au niveau de sous famille associée à une méthode efficace : la PCR compétitive allèle spécifique (la technologie KASPar) ouvre la voie pour la mise en œuvre d'une clé taxonomique moléculaire très efficace et à faible coût.

Pour les groupes récemment divergés comme celui des agrumes vrais qui inclus les formes cultivées, la publication récente des données des génomes chloroplastiques complets de 34 accessions et leur analyse phylogénétique correspondante (Carbonnel Caballero *et al.*, 2015) atteste que les polymorphismes diagnostiques de différenciation des différents genres et de tous les *Citrus* de Swingle & Recce pourraient être trouvés. Par conséquent, concernant le groupe des agrumes vrais et de point de vue phylogénie maternelle, l'ensemble des marqueurs peut être complété et l'approche développée produira une clé moléculaire cladistique très efficace au niveau interspécifique. Cependant, la phylogénie chloroplastique ne révèle qu'une part de l'histoire évolutive du pool génétique dûe à son hérédité monoparentale et elle n'est pas suffisamment résolutive pour les groupes

très récemment divergés à cause de son faible taux de variation. Pour cette raison, les données chloroplastiques et nucléaires sont complémentaires dans la clarification de la structure phylogénomique et de l'histoire évolutive des groupes récemment divergés particulièrement quand les événements de réticulation sont fréquents. Ce qui est le cas des *Citrus* cultivés développés à partir d'hybridation entre taxons ancestraux à mélange nucléaire interspécifique (Garcia-Lor *et al.*, 2013; Curk *et al.*, 2014). Les marqueurs nucléaires diagnostiques des quatre taxons de base du genre *Citrus* couplés à des marqueurs cytoplasmiques sont puissants pour décrypter l'origine des espèces secondaires (Curk *et al.*, 2015; 2016) et les données issues du séquençage complet des génomes produisent une image détaillée de la mosaïque interspécifique nucléaire de certains cultivars (Wu *et al.*, 2014). Le séquençage de nouvelle génération couplé à la réduction de la complexité du génome en utilisant la méthode du Génomage par Séquençage (GBS) ou Génomage des séquences d'ADN associées aux sites de restriction (RAD- seq : *Restriction site Associated DNA Sequencing*) apparaît comme une option convenable pour produire une image phylogénomique sur l'ensemble du génome nucléaire complémentaire des SNPs diagnostiques chloroplastiques.

4. Conclusion

L'analyse de l'ADN chloroplastique est une approche puissante dans les études phylogénétiques à un niveau taxonomique élevé. L'hérédité monoparentale est particulièrement utile dans les pools génétiques à évolution réticulée tel que celle des agrumes cultivés permettant ainsi une déduction aisée de l'ascendance maternelle. Plusieurs méthodes ont été développées pour analyser l'ADN chloroplastique tel que les marqueurs CAPS, SSRs, les amplicons séquencés et plus récemment l'utilisation des données des séquences des génomes complets mappées sur le génome chloroplastique de référence. Dans la présente étude et pour la première fois nous proposons l'utilisation des SNPs diagnostiques de différents clades pour établir une clé taxonomique moléculaire de la sous famille des Aurantioideae.

Il est important de rappeler que le concept était de focaliser l'analyse moléculaire dans des polymorphismes ciblés avec une valeur taxonomique élevée et ce pour analyser ces polymorphismes avec une simple, rentable et très robuste méthode de génotypage ayant une large transférabilité sur la sous famille des Aurantioideae. Les SNPs diagnostiques des clades ont été performés *in silico* à partir des séquences chloroplastiques publiquement disponibles en considérant la classification et les clades identifiés par Bayer *et al.* (2009) comme une référence. Par la suite, nous avons testé l'efficacité de la méthode PCR compétitive allèle spécifique (KASPar) pour atteindre nos objectifs d'une large transférabilité. A partir de 40 SNPs diagnostiques sélectionnés de l'analyse de l'exploration

in silico, 27 marqueurs KASPar diagnostiques de 24 clades ont été correctement développés avec un taux élevé de transférabilité dans la sous famille des Aurantioideae et avec en moyenne seulement 1.2% de données manquantes.

Vingt-un clés marqueurs de clade catégoriquement diagnostiques de 19 clades ont été identifiées. Les deux marqueurs diagnostiques du clade de *Fortunella* comprenant (clade X) ou non (clade AA) *C. halimi* s'adaptent parfaitement à tous les échantillons de *Fortunella* et de *C. halimi* mais une (*Atalantia ceylanica*) et deux Rutaceae (*Murraya Koenigii* et *Clausena anisata*) en dehors du clade Q partagent les allèles diagnostiques. Ces deux marqueurs demeurent utiles dans la résolution de l'intra-clade Q et de l'organisation des agrumes vrais avec un diagnostique ambigu de *Fortunella* et de *C. halimi* au sein de ce clade. Même si leur utilité diagnostique est assez élevée, la validité des quatre autres marqueurs, théoriquement diagnostiques des clades M, N et P reste à valider sur la base d'une structuration précisée du clade M qui est mise en question par l'absence de SNPs diagnostiques pour certains sous-clades et des valeurs diagnostiques incomplètes des marqueurs développés.

Nous avons démontré qu'avec des clades bien établis, des SNPs diagnostiques peuvent être sélectionnés et transformés efficacement en marqueurs de PCR compétitive allèle spécifique en permettant une analyse cladistique rentable très efficace sur une large collection au niveau de la sous famille. La robustesse de la méthode est un avantage déterminant comparé avec d'autres types de marqueurs moléculaires chloroplastiques pour la collaboration d'organisme de recherche et la transférabilité des données entre laboratoires. L'application est considérablement plus facile (travail expérimental et analyse des données) que la comparaison des séquences ciblées. Pour le genre *Citrus* la disponibilité des données des séquences des génomes complets pour les principales espèces des agrumes vrais permettrait bientôt de développer des marqueurs SNPs diagnostiques permettant une résolution fine.

Chapitre III : Révélation des structures interspécifiques *C. maxima*/*C. reticulata* au niveau des génomes des variétés de mandariniers modernes et des hybrides de mandariniers par la technique du Génotypage par Séquençage (GBS)

1. Recherche des marqueurs diagnostiques de la différenciation de *C. maxima* /*C. reticulata*

La procédure d'identification des marqueurs diagnostiques de la différenciation de *C. reticulata*/*C. maxima* à partir de la matrice SNP/InDel issue des lectures (*reads*) élaborées à partir de la technique GBS est schématisée au niveau de la Figure 41. La première étape consiste en un appel des génotypes qui a été élaboré à l'aide du pipeline informatique. La seconde étape consiste à sélectionner les positions polymorphes avec moins de 70% de données manquantes chez les 17 variétés ancestrales représentatives de *C. reticulata* et de *C. maxima* et présentant une structuration interspécifique avec un G_{ST} supérieur à 0,5. La troisième étape concerne l'identification des régions introgressées au niveau des génomes des variétés représentatives de *C. reticulata* et de *C. maxima*. En effet, il ne s'agit pas de vrais ancêtres mais de variétés actuelles résultantes de processus de domestication.

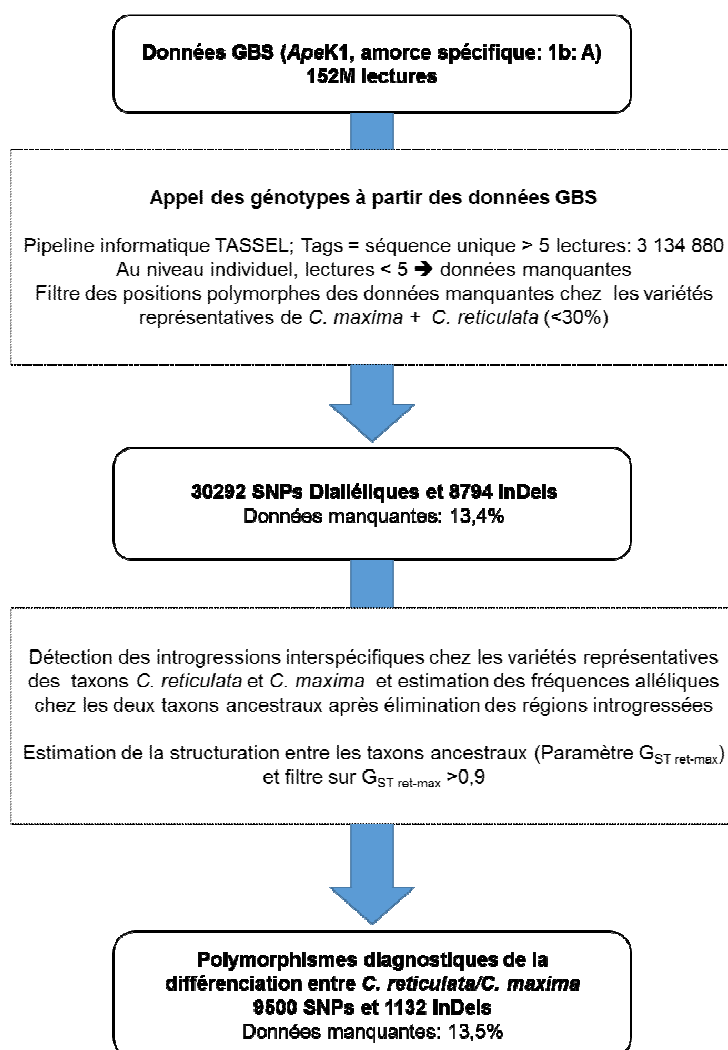


Figure 41 : Procédure d'identification des marqueurs diagnostiques de la différenciation de *C. reticulata*/*C. maxima* à partir des données GBS

De récentes études ont mis en évidence des introgressions interspécifiques chez ces variétés qui étaient antérieurement considérées comme pures (Wu *et al.*, 2014; Curk *et al.*, 2015). Par conséquent, l'identification des SNPs diagnostiques des taxons ancestraux exige une suppression de ces régions introgressées, des variétés concernées, avant d'estimer les fréquences alléliques chez les taxons ancestraux ainsi que le paramètre de différenciation interspécifique ($G_{ST \text{ ret-max}}$). Suite à cette réévaluation des paramètres génétiques, les SNPs et InDels avec $G_{ST \text{ ret-max}} > 0,9$ ont été considérés comme marqueurs diagnostiques.

2. Appel des génotypes et diversité variétale

55 variétés (Annexe 5) ont été séquencées *via* un multiplexage 55 par ligne *Illumina HiSeq 2000* suivant le protocole de Cornell GBS (Elshire *et al.*, 2011) en utilisant l'enzyme de restriction *ApeK1* et une amorce avec une base de sélection ajoutée (A). Cette enzyme de restriction a été retenue chez le maïs et l'orge (Poland *et al.*, 2012). *ApeK1* a été sélectionnée comme la plus robuste enzyme utilisée pour des analyses GBS. En effet, par rapport aux enzymes *PstI* et *MspI*, *ApeK1* a également montrée une distribution intéressante de ses sites de restriction après une digestion *in silico* du génome du soja. Elle a généré une bonne proportion de fragments, assez courts et efficaces pour l'amplification et le séquençage sur la plateforme *Illumina* (Sonah *et al.*, 2013).

Dans notre expérimentation, nous avons obtenu plus de 152 M lectures avec une profondeur suffisante pour un appel de génotype efficace à l'aide du pipeline informatique TASSEL. Ainsi, 91,1% des lectures obtenues ont été validées (code barre : enzyme de restriction plus un insert) et 84,3% ont été positionnées sur le génome de référence de la Clémentine (Wu *et al.*, 2014; *phytozome.net*).

Pour chaque individu, les positions avec moins de cinq lectures ont été considérées comme données manquantes.

Finalement, nous avons sélectionné seulement les positions polymorphes où le taux global de données manquantes chez les variétés représentatives de *C. maxima* et de *C. reticulata* est inférieur à 30%. Un total de 39086 positions polymorphes dont 30292 SNPs et 8794 InDels a été identifié. Le nombre des SNPs varie entre 2581 au niveau du chromosome 6 et 5440 au niveau du chromosome 3. Alors que les InDels varient de 735 pour le chromosome 1 à 1649 pour le chromosome 3 (Tableau 11).

L'indice de fixation de Wright F_{IS} révèle un léger déficit d'hétérozygotie par rapport à l'équilibre de Hardy Wendberg ($F_{IS} = 0,104$ et $0,085$ respectivement pour les SNPs et les InDels) (Tableau 11).

Tableau 11 : Polymorphisme issue des données GBS des 55 variétés étudiées

	SNP				InDel			
	N	Ho	Ha	F _{IS}	N	Ho	Ha	F _{IS}
CH1	2660	0,241 ± 0,006	0,311 ± 0,004	0,230 ± 0,012	735	0,253 ± 0,012	0,313 ± 0,009	0,208 ± 0,025
CH2	3570	0,311 ± 0,005	0,352 ± 0,004	0,117 ± 0,009	947	0,338 ± 0,011	0,365 ± 0,008	0,078 ± 0,018
CH3	5440	0,308 ± 0,005	0,331 ± 0,004	0,087 ± 0,008	1649	0,322 ± 0,009	0,344 ± 0,007	0,087 ± 0,015
CH4	3035	0,287 ± 0,005	0,333 ± 0,005	0,128 ± 0,009	880	0,294 ± 0,012	0,326 ± 0,009	0,121 ± 0,02
CH5	4094	0,298 ± 0,006	0,323 ± 0,005	0,099 ± 0,01	1218	0,294 ± 0,011	0,307 ± 0,009	0,059 ± 0,016
CH6	2581	0,307 ± 0,007	0,336 ± 0,006	0,089 ± 0,01	815	0,332 ± 0,013	0,350 ± 0,01	0,063 ± 0,019
CH7	2740	0,296 ± 0,006	0,318 ± 0,006	0,077 ± 0,01	800	0,321 ± 0,013	0,334 ± 0,01	0,063 ± 0,022
CH8	2719	0,330 ± 0,007	0,341 ± 0,005	0,038 ± 0,01	900	0,351 ± 0,014	0,346 ± 0,01	0,013 ± 0,019
CH9	3453	0,305 ± 0,007	0,320 ± 0,005	0,093 ± 0,012	850	0,313 ± 0,015	0,330 ± 0,009	0,102 ± 0,026
Total	30292	0,300 ± 0,002	0,330 ± 0,002	0,104 ± 0,003	8794	0,314 ± 0,004	0,335 ± 0,003	0,085 ± 0,007

N: nombre de polymorphisme; Ho : Hétérozygotie observée; Ha: Hétérozygotie attendue; F_{IS}: Indice de fixation de Wright

L'hétérozygotie attendue (Ha) est très similaire entre les SNPs (0,330) et les InDels (0,335) et paraît stable chez les différents chromosomes. De même pour l'hétérozygotie observée (Ho) qui enregistre une moyenne de 0,300 et 0,314 respectivement pour les SNPs et les InDels.

3. Diversité entre les mandarines et les pamplemousses et recherche de polymorphismes diagnostiques de la différenciation de *C. reticulata*/*C. maxima*

Toutes les analyses ont été basées sur les positions polymorphes inférieures à 30% de données manquantes chez les variétés représentatives de *C. reticulata*/*C. maxima*.

Le nombre de positions polymorphes est de : 26076, 12682 et 37430 respectivement chez les mandarines, les pamplemousses et les mandarines plus les pamplemousses (Tableau 12 et Tableau 13). En effet, 95,76% (N=37430) du polymorphisme détecté dans l'ensemble des 55 variétés a été identifié chez les 17 variétés représentatives des mandarines et des pamplemousses (Tableau 13). De plus, le nombre des sites polymorphes identifié chez l'ensemble des mandarines est deux fois plus élevé que celui identifié chez les pamplemousses. De même l'hétérozygotie attendue détectée chez les mandarines (Ha= 0,184) est plus élevée que celle enregistrée chez les pamplemousses (Ha= 0,119). Ainsi, les mandarines apparaissent plus polymorphes que les pamplemousses.

Tableau 12 : Polymorphismes (SNP et InDel) identifiés à partir des données GBS chez les 11 variétés de mandarines et les 6 variétés de pamplemousses

	Mandarine				Pamplemousse			
	N	Ho	Ha	F _{IS}	N	Ho	Ha	F _{IS}
CH1	1517	0,167 ± 0,008	0,141 ± 0,006	-0,138 ± 0,017	942	0,129 ± 0,008	0,103 ± 0,006	-0,220 ± 0,027
CH2	3077	0,231 ± 0,007	0,210 ± 0,005	-0,083 ± 0,011	1618	0,151 ± 0,007	0,129 ± 0,005	-0,149 ± 0,022
CH3	5075	0,220 ± 0,006	0,196 ± 0,004	-0,078 ± 0,010	2145	0,132 ± 0,006	0,108 ± 0,004	-0,191 ± 0,018
CH4	1947	0,16 ± 0,007	0,145 ± 0,006	-0,083 ± 0,014	1049	0,111 ± 0,007	0,098 ± 0,005	-0,130 ± 0,026
CH5	3322	0,186 ± 0,007	0,153 ± 0,005	-0,121 ± 0,012	1870	0,165 ± 0,007	0,133 ± 0,005	-0,212 ± 0,021
CH6	2664	0,241 ± 0,007	0,213 ± 0,005	-0,104 ± 0,012	966	0,138 ± 0,009	0,106 ± 0,006	-0,262 ± 0,025
CH7	2663	0,192 ± 0,007	0,177 ± 0,005	-0,063 ± 0,011	1202	0,139 ± 0,008	0,121 ± 0,006	-0,130 ± 0,026
CH8	2877	0,255 ± 0,007	0,232 ± 0,006	-0,074 ± 0,012	1438	0,198 ± 0,010	0,151 ± 0,007	-0,263 ± 0,024
CH9	2934	0,213 ± 0,007	0,190 ± 0,005	-0,099 ± 0,011	1452	0,164 ± 0,009	0,127 ± 0,006	-0,241 ± 0,026
Total	26076	0,208 ± 0,002	0,184 ± 0,002	-0,091 ± 0,004	12682	0,147 ± 0,003	0,119 ± 0,002	-0,199 ± 0,008

N: nombre de polymorphisme; Ho : Hétérozygotie observée; Ha: Hétérozygotie attendue; F_{IS}: Indice de fixation de Wright

D'autre part, les valeurs négatives de l'indice de fixation de Wright reflètent une hétérozygotie excédentaire au sein de chaque groupe (F_{IS}= -0,091 et F_{IS}= -0,199 respectivement chez les mandarines et les pamplemousses). La différenciation entre les deux groupes représentatifs est élevée, comme l'indique les valeurs moyennes de l'indice de fixation de Wright (F_{IS}= 0,349) et du G_{ST} (G_{ST}= 0,471) (Tableau 13).

Tableau 13 : Polymorphismes (SNP et InDel) identifié à partir des données GBS chez les 11 variétés de mandarines plus les 6 variétés de pamplemousses

	Mandarine + Pamplemousse				
	N	Ho	Ha	F _{IS}	G _{ST}
CH1	3395	0,156 ± 0,006	0,392 ± 0,005	0,459 ± 0,019	0,573 ± 0,013
CH2	4517	0,204 ± 0,005	0,373 ± 0,004	0,339 ± 0,015	0,446 ± 0,010
CH3	6670	0,192 ± 0,004	0,355 ± 0,004	0,339 ± 0,013	0,465 ± 0,008
CH4	3686	0,144 ± 0,005	0,358 ± 0,005	0,449 ± 0,018	0,540 ± 0,012
CH5	4975	0,182 ± 0,005	0,350 ± 0,005	0,346 ± 0,016	0,478 ± 0,010
CH6	3239	0,206 ± 0,005	0,366 ± 0,005	0,320 ± 0,017	0,468 ± 0,011
CH7	3342	0,175 ± 0,005	0,351 ± 0,006	0,367 ± 0,017	0,460 ± 0,012
CH8	3448	0,239 ± 0,006	0,351 ± 0,005	0,229 ± 0,016	0,372 ± 0,010
CH9	4158	0,199 ± 0,005	0,360 ± 0,005	0,305 ± 0,017	0,452 ± 0,011
Total	37430	0,189 ± 0,002	0,361 ± 0,002	0,349 ± 0,005	0,471 ± 0,004

N: nombre de polymorphisme; Ho : Hétérozygotie observée; Ha: Hétérozygotie attendue; F_{IS}: Indice de fixation de Wright

En outre, la topologie de l'arbre issu de l'analyse NJ permet de distinguer entre deux groupes majeurs. Elle montre ainsi une différenciation nette entre le groupe des mandarines (coloré en rouge) et celui des pamplemousses (coloré en bleu) et par ailleurs, confirme les résultats obtenus précédemment (Figure 42).

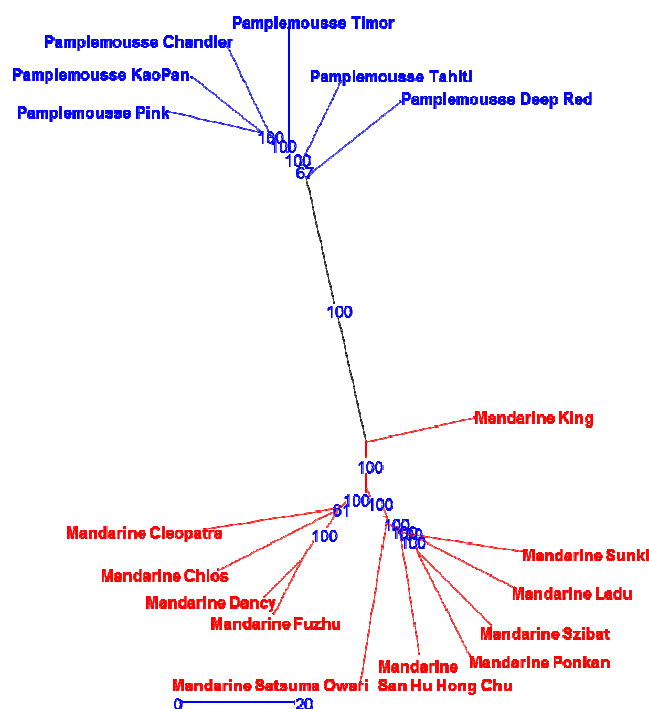


Figure 42 : Arbre issu de l'analyse NJ des variétés représentatives des mandarines et des pamplemousses

4. Introgression interspécifique chez les variétés représentatives des mandarines et des pamplemousses

La recherche des régions introgressées chez les variétés représentatives des taxons de base a été élaborée en ayant recours à deux paramètres principaux analysés sur les neuf chromosomes :

- L'hétérozygotie observée estimée à partir de la matrice des positions SNPs et InDels.
- La similarité des variétés considérées avec le centroïde des variétés représentatives des taxons de base et la moyenne des fréquences alléliques de l'ensemble des variétés représentatives des taxons considérés.

En premier lieu, les SNPs et les InDels avec un $G_{ST} > 1$ (différenciation entre les variétés représentatives des groupes des mandarines et des pamplemousses) supérieur à 0,5 ont été

sélectionnés pour analyser l'hétérozygotie observée et sa similarité avec le centroïde des variétés des mandarines et des pamplemousses. Afin d'identifier les régions d'introgression le long des chromosomes, chaque paramètre a été évalué sur des fenêtres successives de 20 marqueurs. La figure 43 (a) montre un exemple du chromosome 2 du cultivar 'Ponkan'. L'introgression de *C. maxima* chez ce cultivar de mandarinier a été clairement identifiée à la fin du chromosome avec une diminution de la similarité avec le centroïde de la mandarine corrélée à l'augmentation de celle avec celui du pamplemousse.

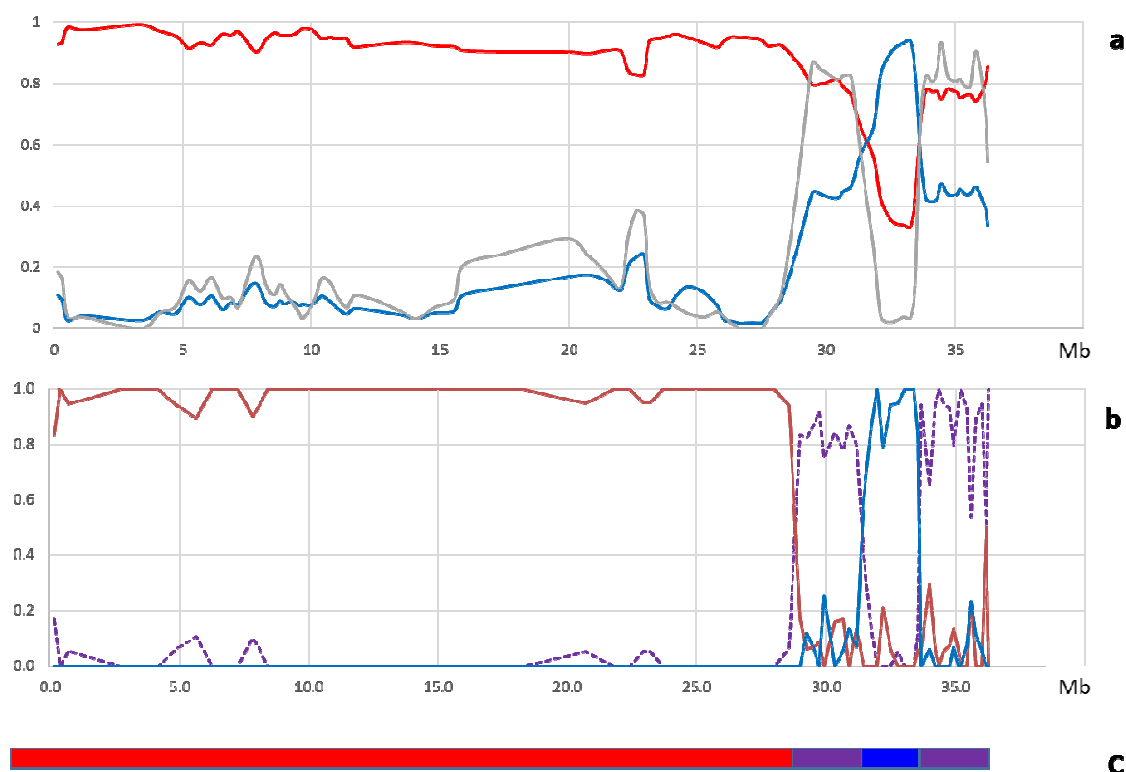


Figure 43 : Identification des introgressions interspécifiques chez les variétés représentatives des taxons ancestraux: exemple du chromosome 2 de la variété 'Ponkan'

a: Hétérozygotie observée (Gris), similarité avec le centroïde des mandarines (Rouge) et pamplemousse (Bleu), fenêtre de 20 marqueurs ($G_{ST} > 0,5$)

b: Proportion de polymorphisme diagnostique de *C. reticulata*/*C. maxima* en homozygotie pour *C. reticulata* (Rouge), pour *C. maxima* (Bleu) et l'hétérozygotie interspécifique (violet).

c: Dédution de la structure phylogénomique du chromosome 2 de la mandarine 'Ponkan' (Rouge: homozygotie *C. reticulata*; Bleu: homozygotie *C. maxima*; Violet: hétérozygotie *C. maxima*/*C. reticulata*).

En outre, le modèle combiné entre l'hétérozygotie observée et la similarité avec les centroïdes mandarines et pamplemousses suggèrent une alternance entre: l'hétérozygotie *C. reticulata*/*C. maxima*, l'homozygotie *C. maxima* et l'hétérozygotie *C. reticulata*/*C. maxima*. Toute cette région qui débute de la position 28 503 155,6 Mb jusqu'à la fin du chromosome a été considérée comme introgressée par *C. maxima*. La même analyse a été faite sur les 17

variétés représentatives des mandarines et des pamplemousses pour les neuf chromosomes. Ces introgressions interspécifiques chez les 17 variétés de référence sont détaillées dans l'Annexe 6.

En second lieu, l'estimation de la fréquence allélique chez les taxons ancestraux *C. reticulata* et *C. maxima* et la différenciation ($G_{ST \text{ ret-max}}$) entre ces deux taxons au niveau de chaque position polymorphe ont été effectués en considérant les régions introgressées chez chaque variété représentative de *C. reticulata* et de *C. maxima* comme données manquantes. La distribution des valeurs du G_{ST} inter-taxon avant et après la suppression des régions introgressées est représentée au niveau de la figure 44. Après l'élimination de ces régions introgressées, 12147 positions polymorphes avec des valeurs de $G_{ST} > 0,9$ ont été identifiées alors qu'elles n'étaient estimées qu'à 6867 positions polymorphes à partir des données initiales.

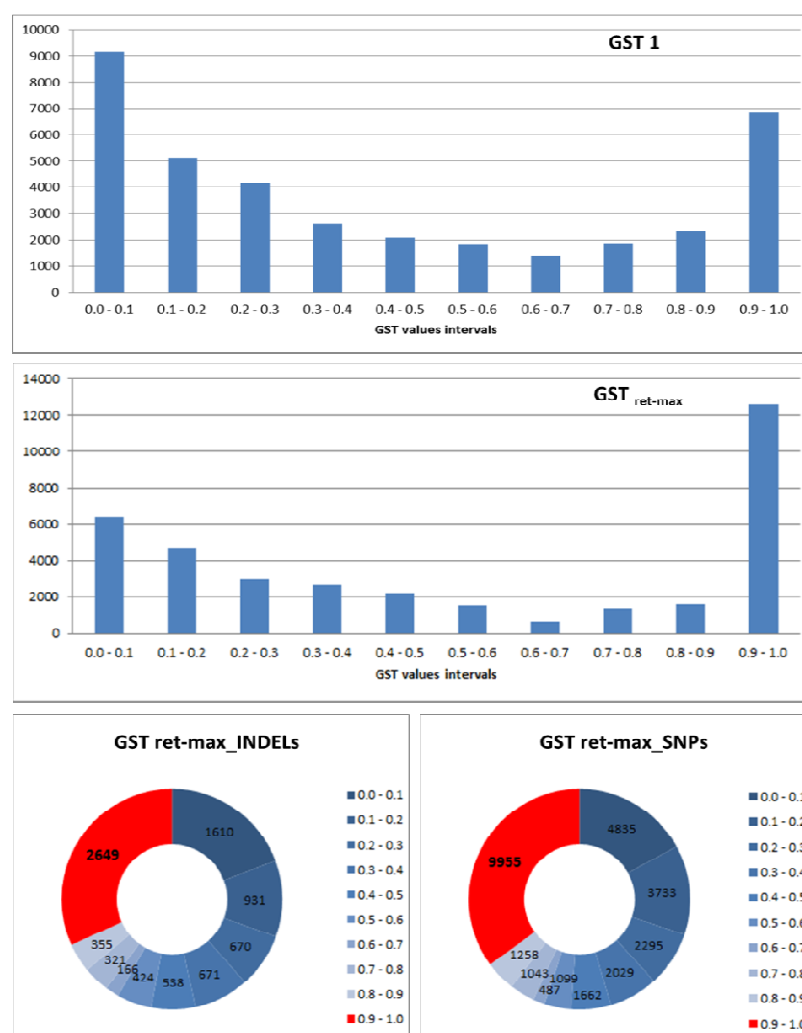


Figure 44 : Comparaison de la distribution du G_{ST} entre les groupes de *C. reticulata* et de *C. maxima* basée sur les données initiales ($G_{ST 1}$) et les données issues après l'élimination des régions introgressées ($G_{ST \text{ ret-max}}$) de type SNP et de type InDel

En troisième lieu, les 12147 positions polymorphes avec des valeurs de G_{ST} supérieures à 0,9 ont été considérées comme des polymorphismes diagnostiques de la différenciation entre *C. reticulata*/*C. maxima* (DPs). Afin d'éviter une information redondante, seule la première position de chaque InDel a été sélectionnée pour les études suivantes sur les structures phylogénomiques des variétés modernes. Ainsi, finalement 10632 DPs (9500 SNPs et 1132 InDels) ont été sélectionnés pour déduire les structures en mosaïques interspécifiques des 55 variétés analysées (Tableau 13).

Tableau 13 : Distribution des DPs entre les neuf chromosomes

	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8	CH9	Total
Total DPs	1234	1392	2127	656	1383	1074	916	814	1036	10632
InDels	125	141	239	65	123	132	106	98	103	1132
SNPs	1109	1251	1888	591	1260	942	810	716	933	9500

Il est intéressant de noter que la distribution des DPs le long des neuf chromosomes est très voisine de celle de l'ensemble des polymorphismes (Figure 45). Ceci témoigne de l'absence de régions privilégiées dans la différenciation entre *C. reticulata* et *C. maxima*. Par ailleurs, les distributions de ces deux classes de polymorphisme suivent assez fidèlement celle des séquences géniques. Ce résultat est cohérent avec les études de coupure *in silico* avec ApeK1 sur la séquence de référence de la Clémentine (P. Ollitrault com. pers.).

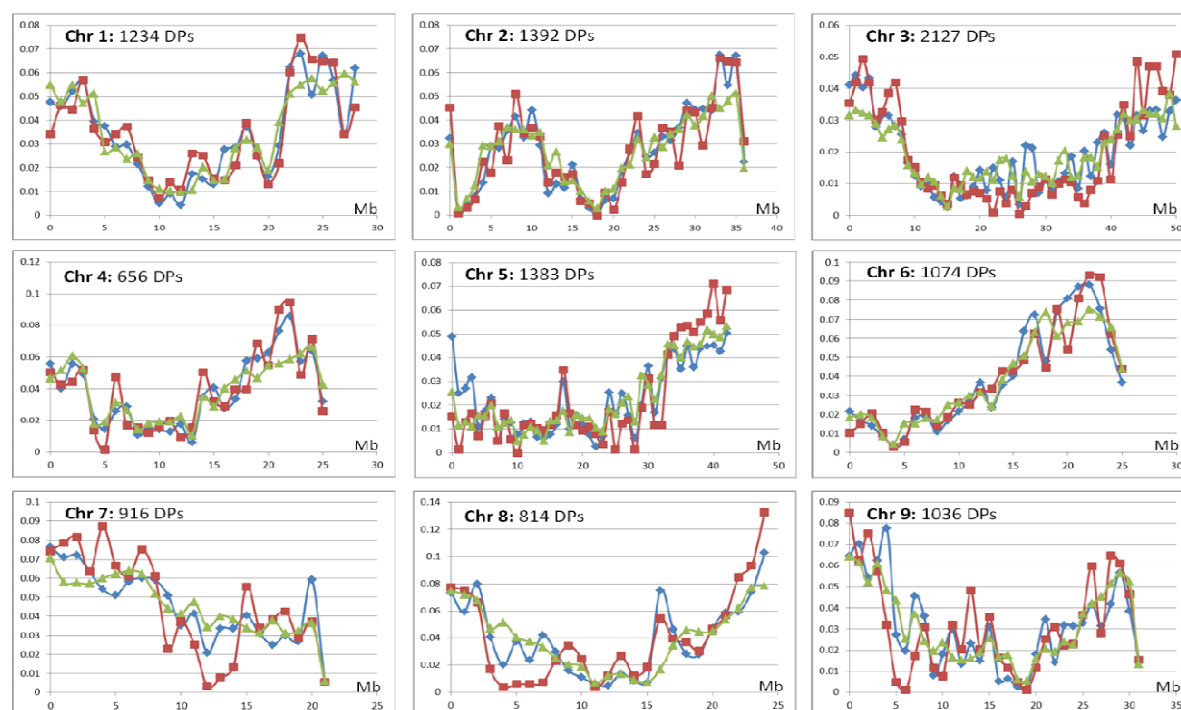


Figure 45 : Comparaison de la distribution des polymorphismes révélés par l'analyse GBS (Bleu), des DPs (Rouge) et des gènes (Vert) le long des neuf chromosomes.

La proportion des marqueurs diagnostiques homozygotes pour *C. reticulata* et *C. maxima* et hétérozygote *C. reticulata/C. maxima* ont été analysées dans des fenêtres successives de 20 marqueurs diagnostiques. Pour 'Ponkan' (Figure 43 b), cette analyse confirme clairement l'hypothèse d'introgression faite à partir des distributions le long du chromosome de l'hétérozygotie et de la similarité avec les centroïdes de *C. reticulata* et de *C. maxima*. La structure en mosaïque interspécifique de 'Ponkan' au niveau du chromosome 2 a été ainsi déduite à partir de cette analyse (Figure 43 c). Cette figure est identique aux conclusions de Wu *et al.* (2014) établies à partir des données de re-séquençage du génome complet.

5. Structure phylogénomique des variétés modernes

Une analyse factorielle a été effectuée afin d'étudier l'organisation de la diversité génétique globale de toutes les variétés analysées. Les 39086 positions polymorphes ont été filtrées pour avoir un maximum de 20% de données manquantes en moyenne sur l'ensemble des variétés. La proportion des données manquantes par chromosome est représentée sur l'Annexe 7. L'orangelo 'Jackson' a été éliminé de cette analyse à cause de sa très haute proportion en données manquantes (68,49%). Par conséquent, l'analyse factorielle a été élaborée sur 54 variétés et 30943 marqueurs. La différenciation pamplemousse/mandarine est l'élément structural principal définissant le premier axe (38,62% de la diversité totale). Le second axe (6,47%) sépare les variétés de mandarines en deux groupes et confirme ainsi l'analyse NJ précédente. L'inertie du troisième et du quatrième axe baisse respectivement à 5,03% et 3,65%. Les groupes des tangors et des tangelos (et les tangelos hybrides) présentent une diversité intra-groupe et n'ont pas été différenciés l'un de l'autre. Ils restent très associés au groupe des mandarines mais présentent globalement un déplacement vers les pamplemousses sur le premier axe. Les oranges amères et les oranges douces se situent sur une position intermédiaire entre les groupes des mandarines et des pamplemousses alors que les pomelos et l'orangelo 'Triumph' sont plus proches du groupe des pamplemousses (Figure 46).

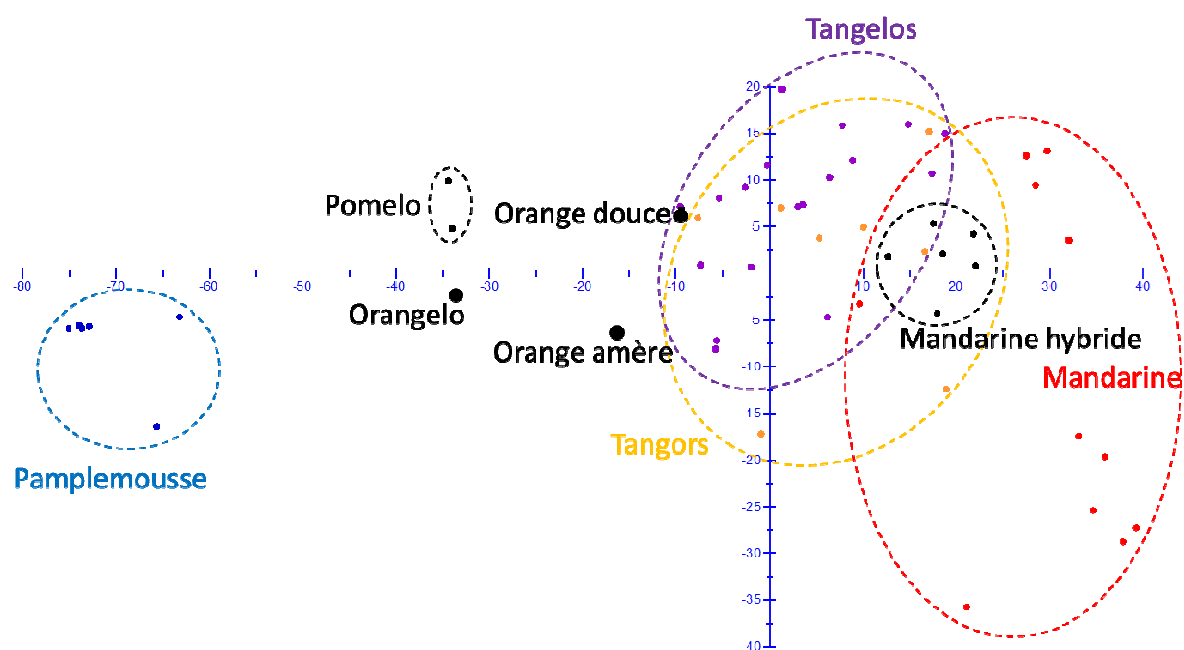


Figure 46 : Analyse factorielle élaborée à partir du tableau de la dissimilarité des 54 variétés sur 30943 marqueurs

La proportion du mélange chez les 55 variétés a été estimée à partir des 10632 positions diagnostiques. Cette figure a été élaborée à partir des données représentées au niveau des Annexes 8 et 9. Les variétés de pamplemoussiers analysées apparaissent être de bons représentants de *C. maxima*. Seules les variétés 'Chandler' et 'Kao Pan' présentent des introgressions de *C. reticulata* toutefois très réduites (0,2%). En revanche, des introgressions de *C. maxima* ont été identifiées chez tous les mandariniers. Toutefois, ces introgressions apparaissent très limitées chez les mandarines 'Cleopatre' et 'Sunki' (<1%) et faible chez les mandarines 'Ladu', 'Szibat' et 'San Hu Hong Chu' (<3%). Des proportions de *C. maxima*, entre 6,5% et 9,1%, ont été trouvées chez les variétés 'Chios', 'Dancy', 'Fuzhu' et 'Ponkan'. Au sein des mandarines considérées comme représentatives de *C. reticulata*, la mandarine 'King' et la mandarine 'Satsuma owari' montrent le taux le plus élevés d'introgression *C. maxima* (respectivement 21,8% et 21,9%). La proportion de *C. maxima* chez les tangors et les tangelos (+ les tangelos hybrides) est comprise respectivement entre 17,7 et 40,8% et entre 17,8 et 40,3 %. La Clémentine a 17,3 % de son génome issu de *C. maxima*. Les deux orangelos et les deux pomelos (*C. paradisi*) présentent des contributions de *C. maxima* entre 63,6% et 65,3%. Les bigaradiers (*C. aurantium*) présentent des contributions similaires de *C. maxima* et *C. reticulata* alors que la contribution de *C. maxima* estimée pour l'orange douce (*C. sinensis*) est de 40,12% (Figure 47).

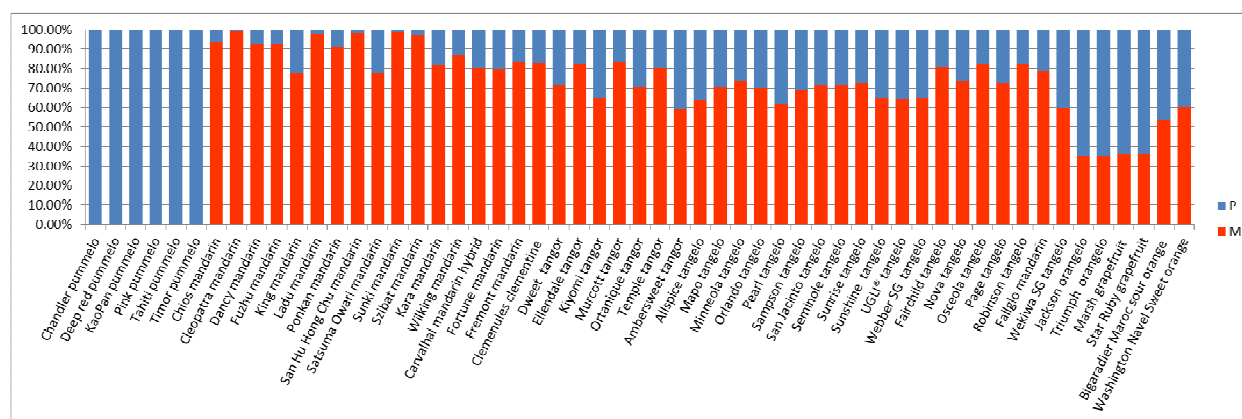


Figure 47 : Proportion globale de *C. maxima* et de *C. reticulata* chez les 55 variétés étudiées estimée à partir des polymorphismes diagnostiques

5.1 Validation de l'approche GBS pour décrypter les structures phylogénomiques par l'analyse de quatre variétés précédemment analysées par WGS (Wu et al. 2014)

La méthode appliquée pour décrypter les structures phylogénomiques est décrite dans la figure 45 de l'analyse du chromosome 2 de l'orange douce 'Washington Navel'. Les paramètres ont été évalués sur une succession de fenêtre de 20 DPs.

La distribution de l'hétérozygotie (Figure 48 a) suggère que deux régions génomiques sont en hétérozygoties interspécifiques et trois fragments en homozygotie phylogénomique. L'étude des marqueurs diagnostiques (Figure 48 b) démontre que la région hétérozygote a une structure interspécifique *C. reticulata*/*C. maxima* alors que les deux premiers fragments homozygotes correspondent à l'homozygote *C. reticulata* et la dernière région homozygote correspond à celle de *C. maxima*. Ainsi, la structure phylogénomique de 'Washington Navel' a été déduite et a été illustrée sur la figure 48 c.

Cette approche a été appliquée à l'ensemble des variétés pour les neuf chromosomes.

L'orange amère 'Bigaradier Maroc' est en hétérozygotie complète entre *C. reticulata* et *C. maxima* (Figure 49). L'origine hybride de l'orange amère entre *C. maxima* et *C. reticulata* a été proposée par des études de marquages moléculaires (Nicolosi et al., 2000 ; Curk et al., 2015), les études de séquences de quelques dizaine de gènes (Garcia Lor et al., 2013) et validée par une étude de séquençage de génome complet (Wu et al., 2014).

L'orange douce 'Washington Navel' montre une structure phylogénomique complexe en hétérozygotie entre *C. reticulata*/*C. maxima* et en homozygotie phylogénomique *C. reticulata* et *C. maxima*. L'homozygotie *C. maxima* a été observé au niveau de la fin du chromosome 2 et au début du chromosome 8 chez 'Washington Navel' (Figure 49). La clémentine 'Clemenules' présentent une mosaïque de régions en homozygotie phylogénomique *C. reticulata* et hétérozygotie *C. maxima*/*C. reticulata* (Figure 49).

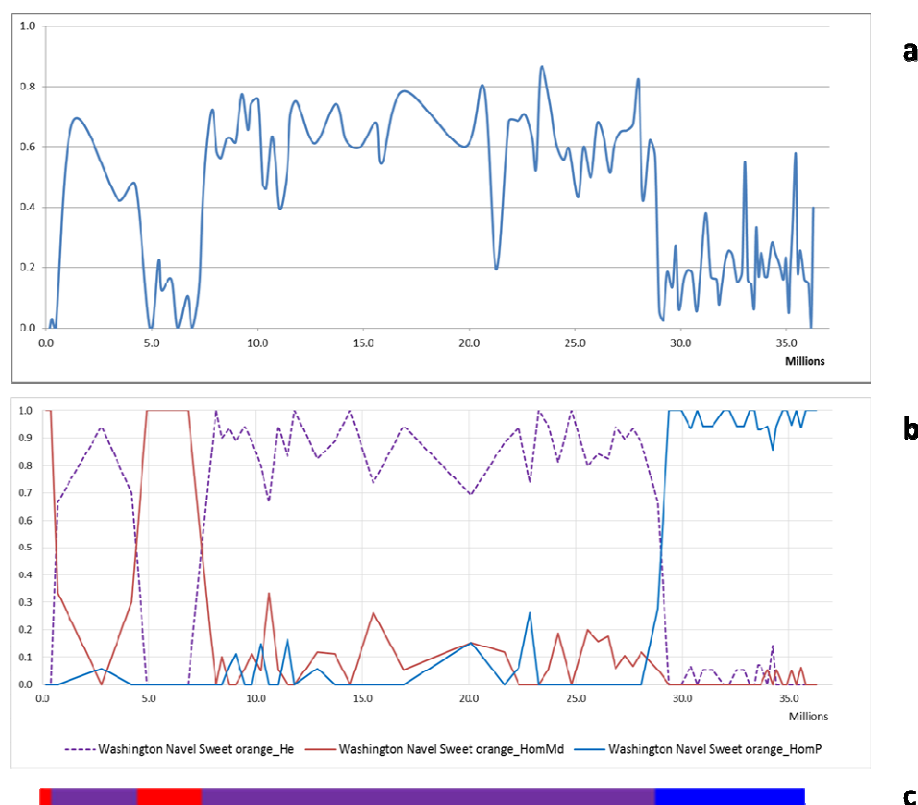


Figure 48 : Déduction de la structure phylogénomique du chromosome 2 de l'orange douce 'Washington Navel'

a : Hétérozygotie

b: Proportion du polymorphisme diagnostique *C. reticulata*/*C. maxima*

c: Déduction de la structure phylogénomique du chromosome 2 de la mandarine 'Ponkan'

Où pour les figures a et b :

■ Homozygote *C. reticulata* , ■ Homozygote *C. maxima*, ■ Hétérozygote interspécifique *C. maxima*/*C. reticulata*

La mandarine 'Ponkan' présente quatre chromosomes (1, 4, 5, 7) en homozygotie phylogénomique *C. reticulata*, quatre chromosomes (3, 6, 8, 9) en mosaïque homozygotie *C. reticulata*, hétérozygotie *C. maxima*/*C. reticulata* et un chromosome (2) en mosaïque avec les deux types d'homozygotie et l'hétérozygotie interspécifique (Figure 51). Ce profil d'introgression *C. maxima* révélé par GBS chez la mandarine 'Ponkan', considérée par les taxonomistes comme une mandarine pure, est en concordance avec les résultats obtenus par l'analyse des séquences du génome complet (Wu *et al.*, 2014).

Des études antérieures (Bretó *et al.*, 2001; Barkley *et al.*, 2006; Ollitrault *et al.*, 2012b; Garcia-Lor *et al.*, 2013b) ont montré que toutes les variétés de l'orange douce sont originaires d'un seul clone de parent via une série de mutations. Curk *et al.* (2014) a obtenu des conclusions similaires pour les variétés de *C. aurantium* montrant une large diversité phénotypique. Ainsi, les résultats obtenus sur 'Washington Navel', 'Clemenules' et l'orange

amère peuvent être extrapolées sur les autres variétés de *C. sinensis*, Clémentine et *C. aurantium*.

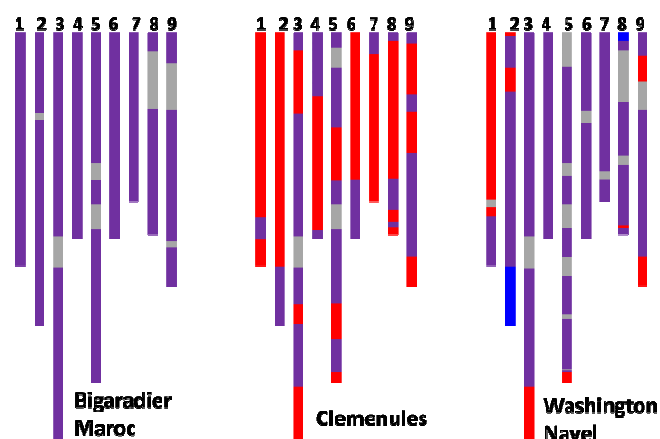


Figure 49 : Caryotype phylogénomique représentatif des neuf chromosomes

■ Homozygote *C. reticulata* , ■ Homozygote *C. maxima*, ■ Hétérozygote Interspécifique *C. maxima/ C. reticulata* et ■ région indéterminée

Les profils de mosaïques interspécifiques déduits à partir des données de GBS pour ces quatre variétés sont totalement concordants avec ceux décrits par Wu *et al.* (2014) à partir de données de séquençage de génomes complets et valident ainsi l'approche GBS et notre procédure de travail.

5.2 Les structures phylogénomiques des variétés représentatives des taxons de base *C. reticulata* et *C. maxima*

Les variétés représentatives des taxons de base pamplemousse : 'Timor', 'Deep Red', 'Pink' et 'Tahiti' semblent être pures *C. maxima* avec aucune introgression *C. reticulata* identifiée. De petites introgressions de *C. reticulata* au niveau du chromosome 2 ont été identifiées pour les pamplemousses 'Chandler' et 'Kao Pan' (Figure 50). Ce résultat confirme l'étude de Wu *et al.* (2014) qui a montré la présence d'un petit fragment (1,5 Mb) d'introgression interspécifique *C. reticulata* chez la variété 'Chandler' au même niveau du chromosome 2.

Ainsi, nos données issues de la technique GBS s'avèrent efficaces dans la déduction de la structure phylogénomique des pamplemoussiers.

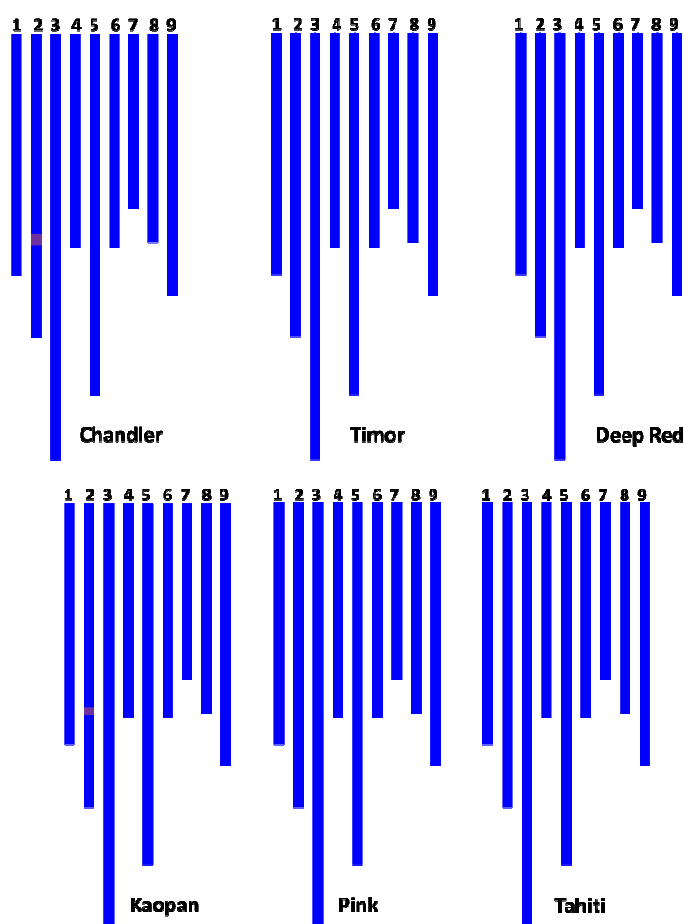


Figure 50 : Caryotype phylogénomique des neuf chromosomes des 6 variétés représentatives du taxon de base *C. maxima*

■ Homozygote *C. reticulata* , ■ Homozygote *C. maxima*, ■ Hétérozygote interspécifique *C. maxima*/*C. reticulata*

Des introgressions de *C. maxima* ont été identifiées chez toutes les variétés de mandarinier. Ainsi, les neuf chromosomes de la variété 'king' ont été trouvés introgressés en hétérozygotie par *C. maxima*.

Sur les neuf chromosomes de la variété 'Satsuma', le chromosome 4 n'est pas introgressé par *C. maxima*. Tandis que le chromosome 8 est introgressé en hétérozygotie et en homozygotie *C. maxima*. Quatre chromosomes introgressés par *C. maxima* ont été observés chez les mandarines 'Fuzhu', 'Dancy' (chromosomes : 2, 3, 6 et 8) et 'Chios' (chromosomes : 3, 4, 6 et 8). Alors que trois chromosomes ont été trouvés avec des introgressions *C. maxima* chez les mandarines 'Szibat' (chromosomes : 2, 4 et 8), 'San Hu Hong Chu' (chromosomes : 2, 8 et 9) et 'Ladu' (chromosomes : 2, 3 et 8) (Figure 51).

Une petite introgression de *C. maxima* chez les variétés 'Cleopatra' et 'Sunki' a été trouvée au niveau du chromosome 3 (Figure 51). Les mandarines 'Cleopatra' et 'Sunki' apparaissent comme les variétés à génomes les plus purs. Ce résultat confirme des déductions antérieures fondées sur une étude d'analyse de structure basée sur les

marqueurs SNP (Curk *et al.*, 2015) et par la méthode de la phylogénie haplotypique de 16 fragments de gènes du chromosome 2 (Curk *et al.*, 2014).

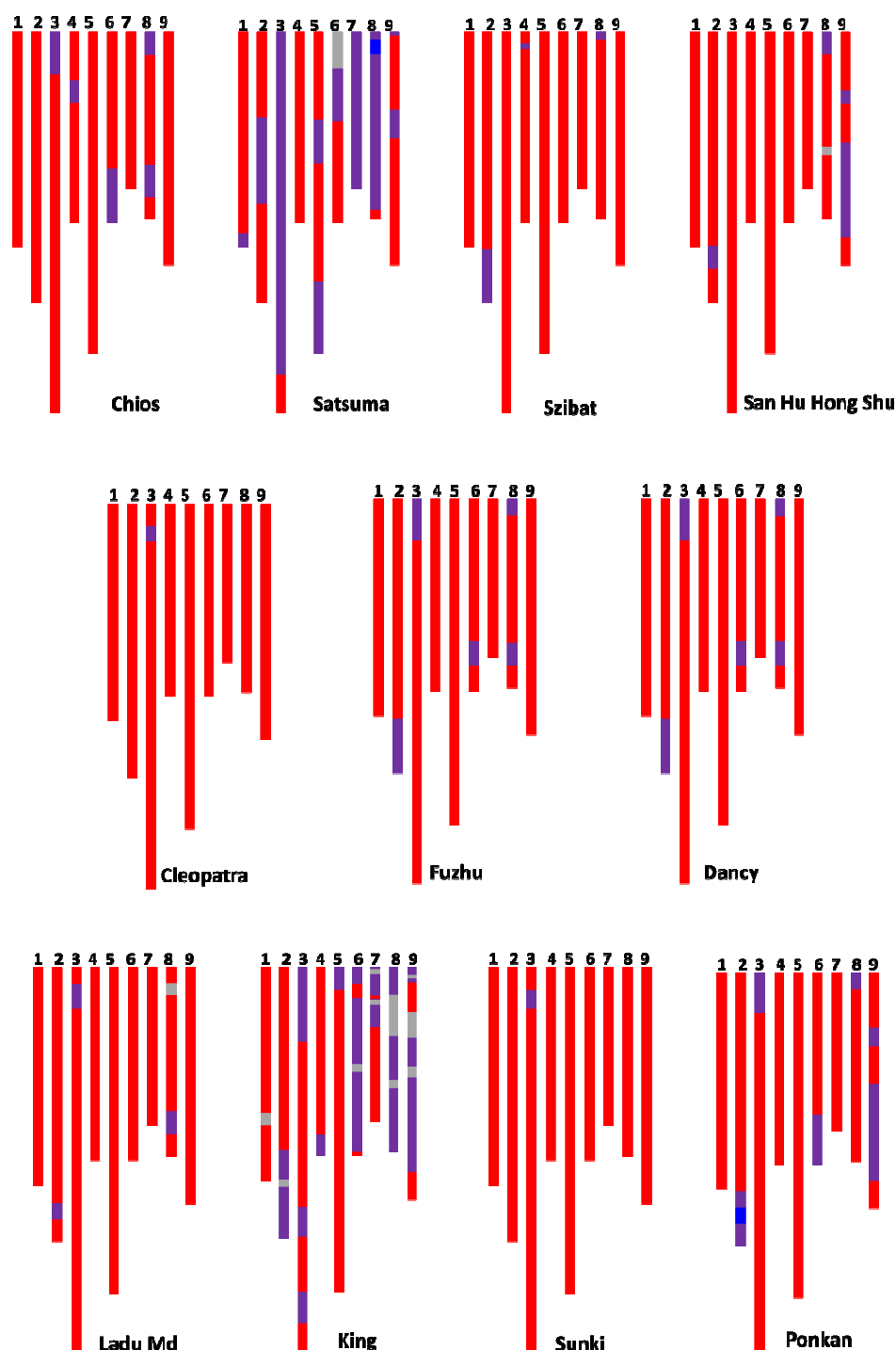


Figure 51 : Caryotype phylogénomique des neuf chromosomes des 11 variétés représentatives du taxon de base *C. reticulata*

■ Homozygote *C. reticulata* , ■ Homozygote *C. maxima*, ■ Hétérozygote interspécifique *C. maxima/ C. reticulata* et ■ région indéterminée

Une étude récente sur une analyse de la structuration du germoplasme des mandarines supportent nos résultats et montrent que le groupe horticole des mandariniers est extrêmement polymorphes. Ainsi, plusieurs mandarines généralement considérées pures montrent des introgressions pamplemousse dans leurs génomes (Garcia-Lor *et al.*, 2015). Nos résultats confirment, de même, les résultats obtenus par Wu *et al.* (2014). Les introgressions *C. maxima* jouent probablement un rôle prépondérant dans la diversité phénotypique des mandariniers cultivées et sont probablement impliqué dans le processus de domestication des mandariniers (Curk *et al.*, 2015).

5.3 Les structures phylogénomiques des mandarines hybrides, tangors, tangelos, pomelos et orangelos

Les structures phylogénomiques des mandarines hybrides, tangors, tangelos, pomelos et orangelos ont été déduites en suivant le modèle décrit au niveau de la figure 43.

Mandarine hybride : Toutes les mandarines hybrides représentent des structures en mosaïque interspécifiques *C. reticulata*/*C. maxima* avec un fond génétique *C. reticulata* prépondérant. Les variétés 'Fallgo' et 'Fortune' et représentent des introgressions interspécifiques *C. maxima*/*C. reticulata* pour les neuf chromosomes. Des chromosomes sans introgressions *C. maxima* (totalement homozygotes *C. reticulata*) ont été identifiés chez 'Carvalhal' (chromosomes 1 et 7), 'Kara' (chromosomes 1 et 4) 'Wilking' (chromosomes 4, 5 et 7) et 'Fremont' (chromosomes 4 et 7)(Figure 52). De l'homozygotie *C. maxima* a été également identifiée au début des chromosomes 3 et 8 de 'Carvalhal', la fin du chromosome 2 et au début des chromosomes 3 et 8 de 'Fallgo', vers le début du chromosome 3 de 'Wilking', à la fin du chromosome 4 et vers la fin du chromosome 6 de 'Fortune' (Figure 52).

Ainsi, ces dernières présentent une structure phylogénomique complexe en hétérozygotie entre *C. reticulata*/*C. maxima* et en homozygotie phylogénomique *C. reticulata* et *C. maxima*. Alors que 'Kara' et 'Fremont' sont en hétérozygotie *C. reticulata*/*C. maxima* et en homozygotie phylogénomique *C. reticulata*.

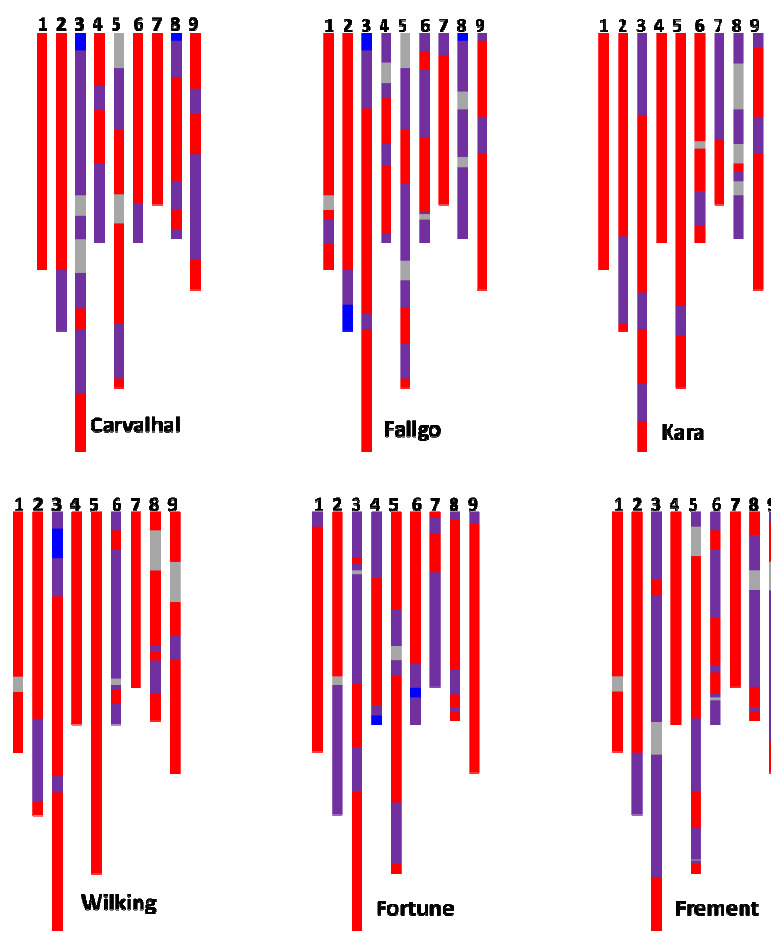


Figure 52 : Caryotype phylogénomique des neuf chromosomes des mandarines hybrides

■ Homozygote *C. reticulata* , ■ Homozygote *C. maxima*, ■ Hétérozygote interspécifique *C. maxima/ C. reticulata* et ■ région indéterminée

Tangor (*C. reticulata* x *C. sinensis*) : Les tangors, les tangelos, les pomelos et les orangelos résultent de la combinaison des pools géniques *C. reticulata* et *C. maxima* avec une contribution variable de chaque taxon. L'analyse de marqueurs diagnostiques des taxons ancestraux a conduit Curk *et al.* (2015) à conclure à une contribution prépondérante de *C. reticulata*.

Les tangors sont très hétérozygotes. Les variétés 'Dweet', 'Kiyomi', 'Ortanique', 'Temple' et 'Ambersweet' représentent une hétérozygotie interspécifique *C. reticulata/ C. maxima* au niveau des neuf chromosomes. Alors que pour 'Ellendale', sur les neuf chromosomes sept présentent des régions en hétérozygotie interspécifique *C. reticulata/ C. maxima* et les deux chromosomes 4 et 6 sont homozygotes *C. reticulata*. Pour la variété 'Murcott', sept chromosomes représentent des zones à hétérozygotie interspécifique *C. reticulata/ C. maxima* et seulement le chromosome 7 présente une homozygotie *C. reticulata* (Figure 53).

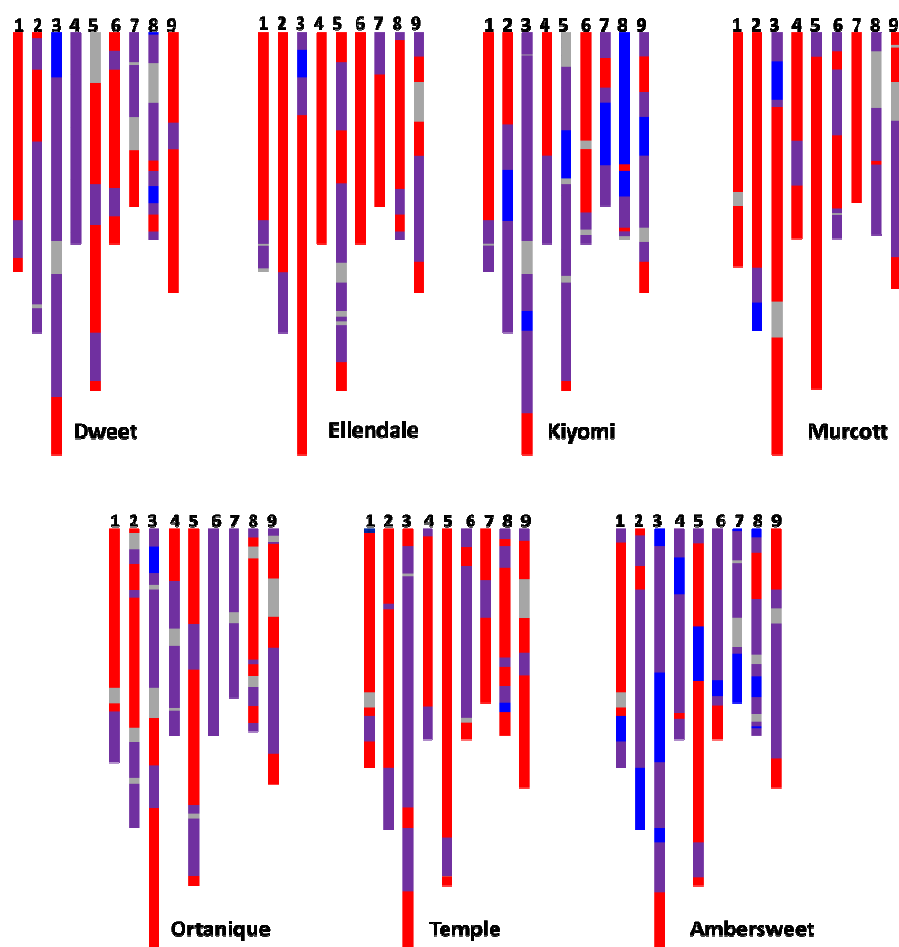


Figure 53 : Caryotype phylogénomique des neuf chromosomes des 7 tangors

■ Homozygote *C. reticulata*, ■ Homozygote *C. maxima*, ■ Hétérozygote interspécifique *C. maxima*/*C. reticulata* et ■ région indéterminée

Des régions homozygotes *C. maxima* ont été identifiées chez les sept tangors analysés et organisées comme suit : au début du chromosome 3, début et vers la fin du chromosome 8 de 'Dweet', au milieu du chromosome 2, vers la fin du chromosome 3, presque au milieu du chromosome 5, vers la fin du chromosome 7, presque tout le chromosome 8 et au milieu du chromosome 9 de 'Kiyomi', vers le début du chromosome 3 de la variété 'Ellendale', la fin du chromosome 2 et presque au début du chromosome 3 de 'Murcott', vers le début du chromosome 3 de la variété 'Ortanique', vers la fin du chromosome 8 de 'Temple', vers la fin du chromosome 1, la fin du chromosome 2, dans plusieurs régions du chromosome 3, vers le début du chromosome 4, au milieu du chromosome 5, vers la fin du chromosome 6, au début et à la fin du chromosome 7, au début et vers la fin du chromosome 8 de la variété 'Ambersweet' (Figure 53). Les tangors Amberswet et Kiyomi présentent une contribution équilibrée de *C. reticulata* et *C. maxima*.

Les tangors présentent ainsi des structures phylogénomiques complexes en hétérozygotie entre *C. reticulata*/ *C. maxima* et en homozygotie phylogénomique *C. reticulata* et *C. maxima*. La contribution prédominante de *C. reticulata* dans la constitution génomique des tangors est confirmée pour cinq d'entre eux (Curk *et al.*, 2015).

Tangelo (*C. reticulata* x *C. paradisi*) : L'hétérozygotie interspécifique *C. reticulata*/ *C. maxima* est clairement identifiée chez toutes les variétés tangelos. En effet, Tous les chromosomes des tangelos représentent de l'hétérozygotie interspécifique *C. reticulata*/ *C. maxima* sauf chez le chromosome 1 des variétés 'Minneola' 'Pearl' et 'Sampson' (Figure 54). La tangelo 'Minneola' est issue d'un rétrocroisement entre un pomelo et une mandarine. Selon une étude basée sur la production d'auraptène et de furocoumarines, la tangelo 'Minneola' s'est avérée ne produire qu'une très faible quantité d'auraptène et aucun furocoumarines. Son chémotype se rapproche ainsi plus du mandarinier que du pomelo (Dugrand-Judek *et al.*, 2015). De même, Curk (2014) a montré que le génotype de 'Minneola' est représenté par une forte contribution de la part de la mandarine (70 %) en comparaison avec le pamplemousse (30 %). Notre structure phylogénomique de 'Minneola' confirme ce résultat et montre une homozygotie *C. reticulata* élevée au niveau de ses différents chromosomes (chromosomes : 1, 3, 4, 5 et 6) (Figure 54).

Des régions du génome en homozygotie *C. maxima* ont été également identifiées chez différentes tangelos. Chez les variétés 'Mappo' et 'Allspice' de l'homozygotie *C. maxima* a été observée au début des chromosomes 3 et 8. Aussi, à la fin du chromosome 2 et vers le début du chromosome 3 de la variété 'Orlando', vers le début du chromosome 3 et vers le début et la fin du chromosome 8 de 'Pearl', à la fin du chromosome 2 et au début du chromosome 8 de 'Sampson', au début du chromosome 3 et 8, vers la fin du chromosome 6 de 'San Jacinto', vers le début du chromosome 3 et vers la fin du chromosome 8 de 'Seminole', à la fin des chromosomes 1 et 2, vers le début du chromosome 3, à la fin du chromosome 5 ainsi qu'au début et à la fin du chromosome 8 de 'Sunrise', à la fin du chromosome 2, vers le début du chromosome 3, au début et vers la fin du chromosome 8 de 'Sunshine', presque la totalité du chromosome 8 de la variété 'Ugli', la fin du chromosome 2, le début et vers la fin du chromosome 8 de la tangelo 'Webber' (Figure 54).

La tangelo 'Minneola' présente une structure phylogénomique en hétérozygotie interspécifique *C. reticulata*/ *C. maxima* et en homozygotie *C. reticulata* et le reste des variétés de tangelo présentent en plus de ça de l'homozygotie *C. maxima*. Les structures phylogénomiques des tangelos étudiées montrent une contribution prépondérante de *C. reticulata* par rapport à *C. maxima*.

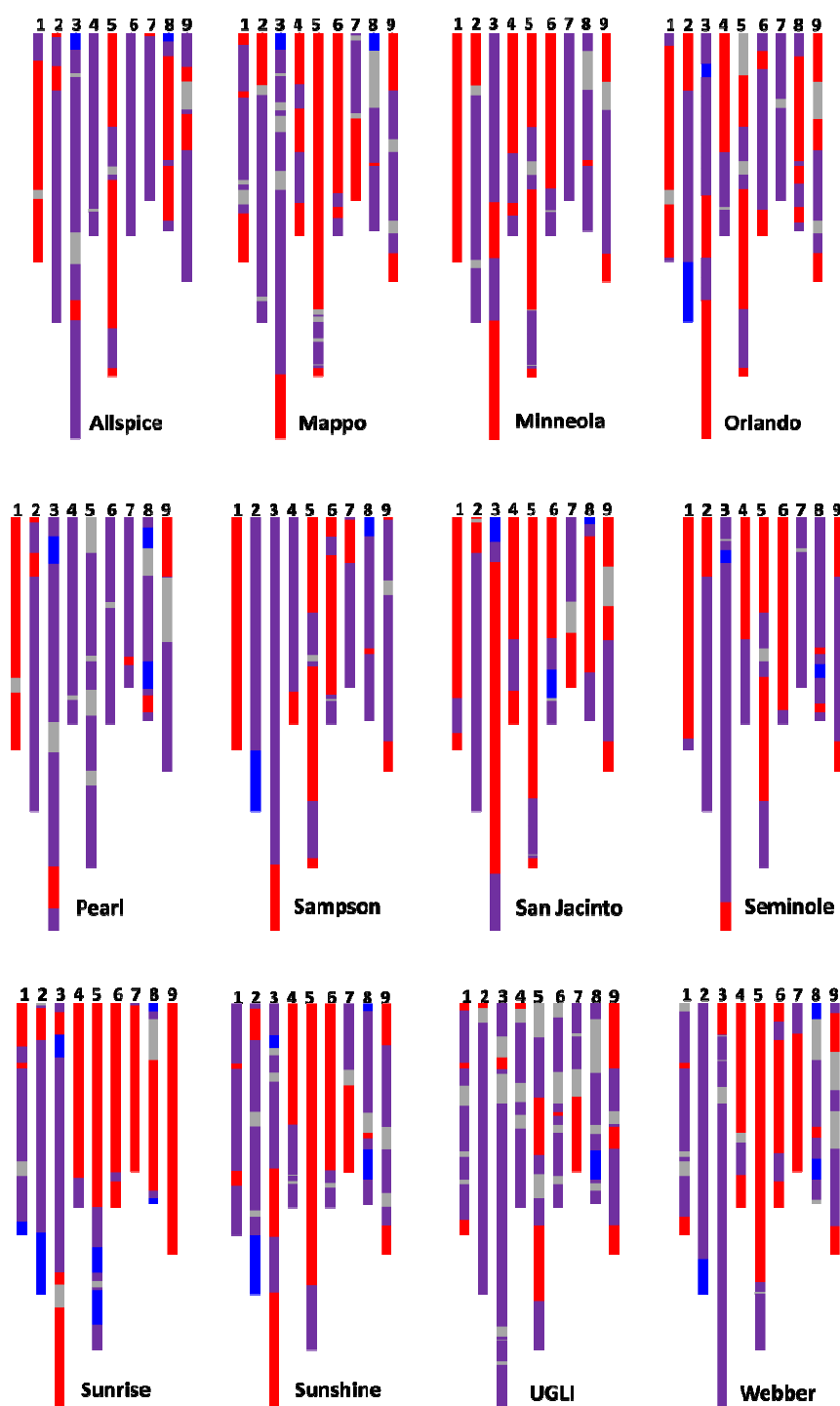


Figure 54 : Caryotype phylogénomique des neuf chromosomes des 12 tangelos

■ Homozygote *C. reticulata* , ■ Homozygote *C. maxima*, ■ Hétérozygote interspécifique *C. maxima*/*C. reticulata* et ■ région Indéterminée

Tangelo hybride (tangor x tangelo) : Toutes les tangelos hybrides présentent de l'hétérozygotie interspécifique *C. reticulata*/*C. maxima* et de l'homozygotie *C. maxima*, *C. reticulata*. Des chromosomes totalement en homozygotie *C. reticulata* ont été observés chez le chromosome 1 de la tangelos 'Page' et les chromosomes 6 et 9 de la tangelo 'Robinson' (Figure 55).

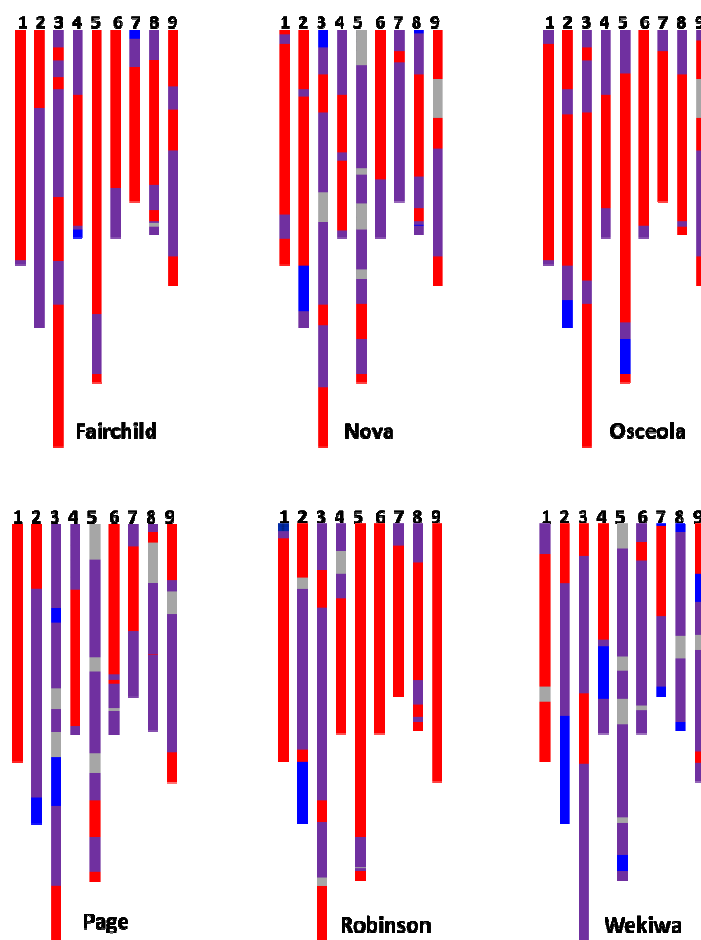


Figure 55 : Caryotype phylogénomique des neuf chromosomes des tangelos hybrides

■ Homozygote *C. reticulata* , ■ Homozygote *C. maxima*, ■ Hétérozygote interspécifique *C. maxima*/*C. reticulata* et ■ région indéterminée

Des régions en homozygotie *C. maxima* ont été observées à la fin du chromosome 4 et au début du chromosome 7 de 'Fairchild', vers la fin du chromosome 2 et au début du chromosome 3 et 8 de 'Nova', à la fin du chromosome 2 et à différentes régions du chromosome 3 de 'Page', à la fin du chromosome 2 de 'Robinson', à la fin du chromosome 2, vers la fin des chromosomes 4 et 5, au début et à la fin des chromosomes 7 et 8 et au milieu du chromosome 9 de 'Wekiwa', à la fin du chromosome 2 et vers la fin du chromosome 5 de 'Osceola' (Figure 55).

Pomelo (*C. paradisi*): Les pomelos montrent une structure phylogénomique hautement hétérozygote interspécifique *C. reticulata*/*C. maxima*, très riches en homozygotie *C. maxima* et avec une homozygotie *C. reticulata* quasiment absente (Figure 56 a). L'hétérozygotie interspécifique *C. reticulata*/*C. maxima* a été observée chez les neuf chromosomes des pomelos 'Marsh' et 'Star Ruby'. Les deux pomelos ont montré une structure phylogénomique très proche voire même identique avec de longs fragments en homozygotie *C. maxima* au niveau des chromosomes 2, 3 et 4, la fin du chromosome 5, le début des chromosomes 7 et 8 et au niveau du chromosome 9. Le chromosome 2 apparaît avec une petite région hétérozygote *C. reticulata*/*C. maxima* suivie jusqu'à la fin du chromosome d'une longue région homozygote *C. maxima* (Figure 56 a). Ce résultat est en cohérence avec l'analyse haplotypique du chromosome 2 de Curk *et al.* (2014).

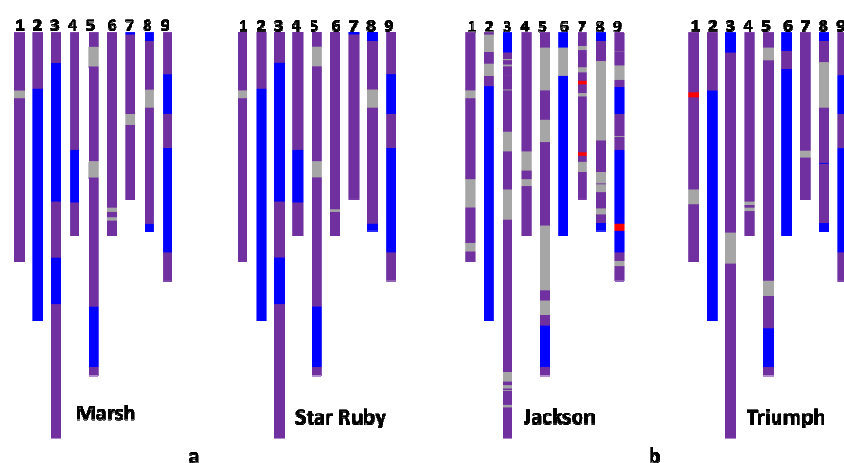


Figure 56 : Caryotype phylogénomique des neuf chromosomes des pomelos et des orangelos

a: Pomelo (*C. paradisi*)

b: Orangelo (*C. sinensis* x *C. paradisi*)

■ Homozygote *C. reticulata* , ■ Homozygote *C. maxima*, ■ Hétérozygote interspécifique *C. maxima*/*C. reticulata* et ■ région indéterminée

Les pomelos proviennent d'une hybridation entre le pamplemousse (*C. maxima*) et l'orange douce (*C. sinensis*). Une étude de la diversité allélique des gènes de biosynthèse des caroténoïdes a montré que le pomelo est en hétérozygotie avec un allèle mandarine et un allèle pamplemousse (Fanciullino *et al.*, 2007). Selon Curk (2014), les pomelos présentent une contribution de 40% *C. reticulata* et 60% *C. maxima*. Nos résultats révélant des régions en hétérozygotie interspécifique *C. reticulata*/*C. maxima* et en homozygotie *C. maxima* confirment ces hypothèses. Nos résultats confirment aussi une étude biochimique qui a montré que chez les pomelos, la dominance (voire une expression transgressive) du pamplemoussier pour la production d'auraptène et d'autres composés est expliquée par la

prédominance de la contribution du pamplemousse dans le génome des pomelos (Dugrand-Judek *et al.*, 2015).

Orangelo (*C. sinensis* x *C. paradisi*) : Les orangelos ‘Jackson’ et ‘Triumph’ présentent des structures phylogénomiques très similaires et complexes. Les régions en hétérozygotie interspécifique *C. reticulata*/*C. maxima* sont prépondérantes. De larges fragments en homozygotie *C. maxima* sont également observés ainsi que quelques rares régions de très petite taille en homozygotie *C. reticulata* (Figure 56 b). L’hétérozygotie interspécifique *C. reticulata*/*C. maxima* a été observée au niveau des neuf chromosomes des orangelos. D’autre part, la même structure phylogénomique des chromosomes 2, 5, 8 et 9 (pour ‘Triumph’) a été observée chez les orangelos et les pomelos. Ces deux variétés sont supposées résulter d’une hybridation naturelle entre *C. sinensis* et *C. paradisi*. Cette étude de la structure phylogénomique est en concordance avec cette hypothèse. Une variation de l’homozygotie *C. maxima* a été observée chez les orangelos par rapport aux pomelos et ce, au début du chromosome 3 et la majorité du chromosome 6 (Figure 56).

De petits fragments en homozygotie *C. reticulata* ont été identifiés chez ‘Jackson’ au niveau des chromosomes 7 et 9 et chez ‘Triumph’ au niveau du chromosome 1 (Figure 56 b). La profondeur de lecture relativement faible pour ces deux orangelos et les nombreuses données manquantes en résultant conduisent à de nombreuses zones d’incertitudes et auraient pu également générer des données artéfactuelles générant l’inférence de ces petits fragments en homozygotie *C. reticulata*. De nouvelles études seront nécessaires pour décrypter finement ces deux variétés.

6. Conclusion

Le groupe horticole des mandariniers est l’un des plus importants agrumes commerciaux. Ce groupe extrêmement polymorphe révèle une large variabilité phénotypique et génétique reflétant sa longue histoire de culture ponctuée par de nombreuses hybridations et mutations naturelles. Les études de marquages moléculaires récentes ont suggéré que toutes les mandarines cultivées présentes des introgressions *C. maxima* dans un fond génétique principalement *C. reticulata*. *C. reticulata* et *C. maxima* sont aussi les ancêtres des oranges douces et amères, des pomelos et des petits agrumes (mandarine hybride, tangelo et tangor). Tirant partie de l’avènement des NGS, nous avons conduit une analyse de la structure phylogénomique sur les neuf chromosomes des variétés modernes de mandarinier, de pamplemoussier et des espèces et groupes horticoles qui en découlent. Cette étude s’est appuyée sur les polymorphismes SNPs et InDels identifiés par une analyse GBS. L’analyse GSB nous a permis d’obtenir une matrice SNP/InDel après une succession d’étapes d’analyse des données grâce au pipeline informatique *Tassel* en utilisant le génome de

référence du Clémentinier. 30292 marqueurs SNPs et 8794 InDels ont été identifiés. La diversité observée chez les mandariniers est supérieure à celle des pamplemoussiers. Une différenciation nette entre les deux groupes horticoles a été révélée par valeur moyenne du G_{ST} (0,471) et de la topologie de la classification arborée (NJ).

Nous avons développé un pipeline d'analyse de donnée efficace pour décrypter les structures phylogéniques des 55 variétés étudiées. Afin d'identifier les polymorphismes diagnostiques de la différenciation *C. reticulata*/*C. maxima* une première étape a consisté en l'identification des régions introgressées au niveau des génomes des variétés représentatives de *C. reticulata* et de *C. maxima*. Elle a été réalisée sur la base des distribution d'hétérozygotie et de similarité par rapport au centroïdes des mandariniers et des pamplemoussiers en se basant sur les positions montrant des structures interspécifiques avec un G_{ST} supérieur à 0,5. Des introgressions *C. maxima* ont été observées chez toutes les variétés de mandarinier. L'élimination de ces régions introgressées a permis de sélectionner les marqueurs diagnostiques en se basant sur les 12564 positions polymorphes avec des valeurs de $G_{ST \text{ ret-max}}$ supérieures à 0,9 alors qu'elles n'étaient que 6867 à partir des données initiales. Ces positions considérées comme diagnostique de différenciation entre *C. reticulata*/*C. maxima* ont été utilisées pour déduire la structure en mosaïque interspécifique des 55 variétés analysées.

Les profils des mosaïques interspécifiques déduits pour l'orange douce, l'orange amère, le mandarinier 'Ponkan' et la clémentine sont identiques à ceux décrit par Wu *et al.* (2014) et valide, ainsi, notre procédure de travail. L'introgression *C. maxima* a été observée chez toutes les variétés mandarines représentatives du taxon de base *C. reticulata*. Cette introgression est toutefois très limitée chez 'Cleopatra' et 'Sunki' qui apparaissent comme les variétés de référence de *C. reticulata* aux génomes les plus purs. Une petite introgression de *C. reticulata* a été observée au niveau du chromosome 2 des variétés 'Chandler' et 'Kao Pan'. Les variétés Timor', 'Deep Red', 'Pink' et 'Tahiti' peuvent être considérées comme de pur représentant du taxon de base *C. maxima*.

Toutes les mandarines hybrides (sauf 'Kara' et 'Fremont') montrent une structure phylogénomique complexe en hétérozygotie entre *C. reticulata*/*C. maxima* et en homozygotie *C. reticulata* et *C. maxima*. 'Kara' et 'Fremont' présentent une mosaïque de fragments en hétérozygotie *C. reticulata*/*C. maxima* et en homozygotie *C. reticulata*. Le génome *C. reticulata* reste toutefois prédominant chez toutes ces mandarines.

Les tangors, les tangelos (sauf 'Minneola') et les tangelos hybrides montrent une structure phylogénomique complexe en hétérozygotie entre *C. reticulata*/*C. maxima* et en homozygotie phylogénomique *C. reticulata* et *C. maxima*. La tangelo 'Minneola' montre une structure phylogénomique en hétérozygotie interspécifique *C. reticulata*/*C. maxima* et en

homozygotie *C. reticulata*. Les tangors, les tangelos et les tangelos hybrides montrent une contribution dominante *C. reticulata* par rapport à *C. maxima*.

Les pomelos montrent des structures phylogénomiques très similaires hautement hétérozygotes interspécifiques *C. reticulata*/*C. maxima*, très riches en homozygotie *C. maxima* et l'absence d'homozygotie *C. reticulata*. La structure phylogénomique des orangelos ressemble à celle des pomelos et confirme l'hypothèse de l'hérédité parentale pomelo des orangelos.

Nous avons démontré la pertinence de la GBS pour étudier des structures phylogénomiques complexes et développé un pipe-line d'analyse performant pour les complexes d'espèces résultants de deux pool géniques ancestraux.

Discussion et Conclusions générales

&

Perspectives

Les agrumes rêvaient une place prépondérante dans l'économie nationale et internationale. Ainsi, Le groupe des agrumes vrais appartenant à la sous famille des Aurantioideae compte six genres interféquents: *Poncirus*, *Fortunella*, *Microcitrus*, *Clymenia*, *Eremocitrus* et *Citrus*. Le genre *Citrus* comprend les fruits cultivés et comestibles. Il comprend quatre espèces de base *C. reticulata* (mandarine), *C. maxima* (pamplemousse), *C. medica* (cédrat) et *C. micrantha* (papeda) à partir desquelles les autres espèces secondaires (orange douce, orange amère, clémentine, lime, citron, pomelo) ont dérivé par des hybridations interspécifiques. Ces espèces sont caractérisées par une taxonomie controversée due à leur origine interspécifique hautement hétérozygote. En raison de diverses caractéristiques biologiques tels que : l'apomixie, l'incompatibilité sexuelle, la stérilité et les structures génomiques d'origine interspécifique hautement hétérozygotes l'amélioration variétale est limitée. Pour ce faire, une bonne connaissance du germoplasme d'agrumes et de son origine génétique permet d'étudier rationnellement leur phylogénie par un décryptage des structures génomiques des variétés en se basant aux pools génétiques des taxons de référence. Ce qui permet une bonne gestion du germoplasme et de développer par ailleurs une stratégie d'amélioration variétale.

C'est dans ce cadre d'étude que la présente thèse a été élaborée en suivant différentes approches d'analyse qui ont combiné les analyses des séquences chloroplastiques, les marqueurs SNPs et KASPar chloroplastiques ainsi que les marqueurs SNPs et InDel nucléaires.

Au niveau du premier chapitre, nous nous sommes intéressés à une collection de vingt-sept variétés appartenant à différentes espèces du genre *Citrus* tunisien afin d'étudier leur diversité génétique cytoplasmique en utilisant l'espaceur intergénique *trnL-trnF*. L'efficacité de cette région dans l'étude de la phylogénie maternelle a guidé notre choix. L'amplification de cette région chloroplastique a généré des amplimères d'une taille approximative de 400 pb. Un séquençage Sanger a été préalablement appliqué à ces amplimères et a montré une variation de la taille et de la teneur en nucléotides entre les différentes séquences. En revanche, nos séquences ont montré une faible variation de la diversité nucléotidique et haplotypique.

D'autre part, l'espaceur intergénique *trnL-trnF* a montré la présence d'une seule copie de pseudogène originale du gène *trnF* chez les vingt-sept variétés analysées au niveau de la position 275 pb avec une taille qui varie de 49 à 63 pb. Le domaine anticodon a été identifié comme étant l'élément le plus conservé, mais une transversion (T→C) a été identifiée au niveau du domaine D. Entre autre, deux transversions (T→A et T→G) ont été identifiées au niveau du domaine D.

Les tests de neutralité (Tajima, Fu & Li et Fu) montrent des valeurs positives et non significatives et plaident, ainsi, en faveur d'un modèle d'évolution neutre et supportent l'hypothèse d'une population de taille constante.

L'arbre phylogénétique de type NJ a identifié l'origine maternelle des espèces du genre *Citrus* étudiées et a montré une séparation nette entre les espèces de base : *C. maxima*, *C. medica* et *C. reticulata* ainsi qu'une distinction entre deux groupes majeurs : le groupe des mandarines et celui des pamplemousses. Cette analyse a permis d'élucider les liens de parentés entre les différentes espèces du genre *Citrus*. Ainsi, le regroupement des espèces *C. aurantium*, *C. sinensis*, *C. limon*, *C. tangelo*, *C. paradisi* et *C. maxima* suggère que *C. maxima* est probablement leur ancêtre commun. D'autre part, les données chloroplastiques confirment l'origine maternelle *C. reticulata* de la clémentine. Nos résultats montrent clairement la contribution des pools génétiques du pamplemousse et de la mandarine dans le développement des espèces secondaires tunisiennes.

L'implication des agrumes d'autres pays du bassin méditerranéen et du monde est une approche élaborée pour analyser la diversité globale afin de vérifier et d'affiner nos conclusions sur l'histoire taxonomique des agrumes.

Pour ce faire, au niveau du deuxième chapitre, nous nous sommes intéressés à la sous famille des Aurantioideae, qui inclue les espèces cultivées des agrumes afin d'élargir l'étude chloroplastique et développer de nouvelles connaissances. Les analyses moléculaires et particulièrement les données des séquences chloroplastiques de plusieurs régions ont mis en évidence l'organisation phylogénétique de cette sous famille. Cependant, le séquençage et l'analyse de différentes régions chloroplastiques surdifficiles à appliquer sur une large collection ou un large échantillon de populations naturelles. L'objectif de ce travail est d'identifier des SNPs diagnostiques des principaux clades établis à partir des séquences chloroplastiques, connues comme très conservées, et de développer une clé moléculaire taxonomique de la sous famille des Aurantioideae basée sur une méthode simple, robuste et rentable du génotypage des marqueurs SNP.

Les séquences nucléotidiques de huit régions génomiques chloroplastiques publiquement disponible à Genbank ont été comparées pour 79 variétés de la sous famille des Aurantioideae afin de rechercher des SNPs révélant une différenciation taxonomique au niveau des inter-tribus, inter-sous tribus, inter-genres et interspécifiques. Des 166 SNPs diagnostiques identifiés, 40 SNPs diagnostiques de 33 clades. Les marqueurs KASPar ont été sélectionnés à partir des 40 SNPs diagnostiques et ont été utilisés chez 108 variétés de la sous famille des Aurantioideae pour tester leur valeur taxonomique, leur transférabilité au sein de la sous famille des Aurantioideae et de spécifier la phylogénie des espèces du genre

Citrus. Vingt-sept marqueurs diagnostiques de 24 clades ont été validés et ont montré un haut taux de transférabilité au sein de la sous famille des Aurantioideae.

L'arbre de type UPGMA dressé à partir des marqueurs validés a montré une organisation cladistique cohérente avec les analyses génétiques de Bayer *et al.*, (2009). Vingt-un marqueurs clés cladistiques, catégoriquement diagnostiques de 19 clades, ont été identifiés. Les deux marqueurs diagnostiques du clade de *Fortunella* comprenant (clade X) ou non (clade AA) *C. halimi* s'adaptent parfaitement à tous les échantillons de *Fortunella* et de *C. halimi* mais respectivement une (*Atalantia ceylanica*) et deux Rutaceae (*Murraya Koenigii* and *Clausena anisata*) en dehors du clade Q partagent les allèles diagnostiques. Ces deux marqueurs demeurent utiles dans la résolution de l'intra-clade Q et de l'organisation des agrumes vrais avec un diagnostique ambigu de *Fortunella* et de *C. halimi* au sein de ce clade. Même si leur utilité diagnostique est assez élevée, la validité des quatre autres marqueurs, théoriquement diagnostiques des clades M, N et P reste à valider sur la base d'une structuration précisée du clade M qui est mise en question par l'absence de SNP diagnostique pour certains sous-clades et des valeurs diagnostiques incomplètes des marqueurs développés.

En particulier, l'origine monophylétique des genres au sein du groupe des 'agrumes vrais' plus *Oxanthera* a été validé. Toutefois, certaines clarifications demeurent nécessaires concernant l'organisation des autres espèces sauvages de la tribu des Citreae.

Nous avons validé le concept qu'avec des clades bien établis, des SNPs diagnostiques peuvent être sélectionnés et efficacement transformés pour une PCR compétitive allèle spécifique KASPar permettant une analyse cladistique très efficace et rentable dans de grandes collections au niveau de sous famille.

L'analyse de l'ADN chloroplastique est une approche puissante dans les études phylogénétiques. L'hérédité monoparentale est particulièrement utile dans les pools génétiques à évolution réticulée tel que le cas des agrumes cultivés permettant ainsi une déduction aisée de l'ascendance maternelle. Mais l'hérédité monoparentale chloroplastique développe une histoire évolutive incomplète vue qu'elle ne présente qu'une part du pool génétique. Ainsi, l'information chloroplastique couplée à l'information nucléaire s'avèrent complémentaires dans le traçage de l'histoire évolutive, surtout, des groupes récemment divergés, dus à leur faible taux de variabilité, et sous évènement de réticulation comme celui des *Citrus*.

Au niveau du troisième chapitre nous nous sommes intéressés au groupe horticole des mandariniers reconnu phénotypiquement très polymorphe. Il est considéré comme un élément important dans la production des fruits des agrumes. Selon des études moléculaires récentes (Wu *et al.*, 2014) toutes les mandarines cultivées analysées portent des

introgressions de *C. maxima* dans le patrimoine génétique de *C. reticulata*. En plus, *C. maxima* et *C. reticulata* sont les ancêtres des oranges douces et amères, les pomelos et de tout le groupe horticol des petits agrumes (hybride de mandarinier, tangor et tangelo). Récemment, la technologie NGS a modifié profondément la manière d'analyser l'évolution des plantes et des structures génomiques, passant des études phylogénétiques aux études phylogénomiques. Pour ce faire, au niveau de ce chapitre nous avons développé une procédure d'étude phylogénomique basée sur des données SNPs et InDels obtenus par l'outil de la technique GBS. Cette étude nous a permis d'analyser la mosaïque interspécifique tout au long des neuf chromosomes des espèces secondaires et des variétés modernes des agrumes résultant de la combinaison des pools génétiques de *C. reticulata* et de *C. maxima*.

L'analyse GBS consiste à séquencer les 55 variétés via le séquençage *ILLUMINA HISEQ2000* en utilisant l'enzyme de restriction *ApeK1* et une amorce avec une base de sélection (A) pour améliorer la profondeur de lecture. Les polymorphismes diagnostiques de la différenciation *C. maxima*/*C. reticulata* ont été identifiés couvrant l'ensemble du génome. Ces points diagnostiques polymorphes ont été utilisés pour décrypter l'origine phylogénomique des fragments des génomes tout au long des neuf chromosomes des 55 variétés.

La diversité observée chez les mandariniers est plus supérieure à celle des pamplemoussiers. Une différenciation nette entre les deux groupes horticoles a été déduite de la valeur moyenne du G_{ST} (0,471) et de la topologie de l'arbre NJ.

12564 positions polymorphes ont été sélectionnées et ont permis d'analyser les structures en mosaïque interspécifiques tout au long des neuf chromosomes. Nos résultats confirment des analyses récentes basées sur des données des séquences des génomes complets de l'orange douce et amère, la mandarine 'Ponkan' et la clémentine 'Nules'. Les introgressions de *C. maxima* ont été observées chez toutes les mandarines mais s'avèrent limitées chez les mandarines 'Cleopatra' et 'Sunki'.

L'introgression *C. maxima* a été observée chez toutes les variétés mandarines représentatives du taxon de base *C. reticulata*. Cette introgression a été limitée chez 'Cleopatra' et 'Sunki' qui apparaissent comme les variétés de référence à génomes les plus purs. Une petite introgression de *C. reticulata* a été observée au niveau du chromosome 2 des variétés 'Chandler' et 'Kao Pan'. Les variétés Timor', 'Deep Red', 'Pink' et 'Tahiti' sont considérées comme de vraies variétés représentatives du taxon de base *C. maxima*. Les caryotypes phylogénomiques des mandarines hybrides, tangors, tangelos, orangelos et pomelos ont été aussi démontrés.

Toutes les mandarines hybrides, exception faite à 'Kara' et 'Fremont', présentent des structures phylogénomiques complexes en hétérozygotie entre *C. reticulata*/*C. maxima* et en

homozygotie *C. reticulata* et *C. maxima*. 'Kara' et 'Fremont' sont en hétérozygotie *C. reticulata/C. maxima* et en homozygotie *C. reticulata*.

Les tangors, les tangelos et les tangelos hybrides montrent une structure phylogénomique complexe en hétérozygotie entre *C. reticulata/C. maxima* et en homozygotie phylogénomique *C. reticulata* et *C. maxima*. Sauf que la tangelo 'Minneola' a montré une structure phylogénomique en hétérozygotie interspécifique *C. reticulata/C. maxima* et en homozygotie *C. reticulata*. Les tangors, les tangelos et les tangelos hybrides montrent une contribution dominante de *C. reticulata* par rapport à *C. maxima*.

Les pomelos 'Marsh' et 'Star Ruby' présentent des structures phylogénomiques très similaires hautement hétérozygote *C. reticulata/C. maxima* et très riches en homozygotie *C. maxima*. Alors que les orangelos montrent une structure phylogénomique hautement hétérozygote interspécifique *C. reticulata/C. maxima*, très riches en homozygotie *C. maxima* et faiblement homozygote *C. reticulata*. La structure phylogénomique des orangelos ressemble à celle des pomelos et confirme l'hypothèse de l'hérédité parentale pomelo des orangelos.

La technique GBS couplée à la détection des points polymorphes diagnostiques était efficace dans le décryptage des caryotypes phylogénomiques des variétés qui dérivent d'une mosaïque de deux espèces ancestrales. Toutes les mandarines, les hybrides de mandariniers, tangelos, tangors analysés montrent une introgression de *C. maxima* à différentes parties du génome. Le mélange *C. reticulata/C. maxima* doit être l'élément majeur de la variabilité phénotypique élevée de ces germoplasmes.

Les résultats obtenus lors de cette thèse contribuent à clarifier les liens de parenté au sein de la sous famille des Aurantioideae et de décrypter les structures phylogénomiques des espèces secondaires du genre *Citrus* et des variétés hybrides. Ces résultats ouvrent la voie vers une optimisation de nouvelle stratégie d'amélioration variétale basée sur les structures phylogénomiques déduites pour améliorer la gestion des ressources phytogénétiques.

Compte tenu de l'importance des résultats obtenus au cours de ce présent travail, nombreuses voies perspectives sont recommandées :

Elargir l'échantillonnage des variétés tunisiennes répertoriées à travers le pays permettait de mieux valoriser les ressources génétiques des agrumes tunisiens.

Explorer l'ADN mitochondrial pour mieux décrire l'hérédité maternelle et préciser la structuration de la diversité génétique cytoplasmique en combinant les marqueurs mitochondriaux et chloroplastiques chez les agrumes tunisiens pour rationner la gestion des ressources et contribuer à l'amélioration variétale.

Avec, la disponibilité des données des séquences des génomes complets des principales espèces du groupe des agrumes vrais, il serait intéressant de poursuivre cette étude. La sélection d'un ensemble de marqueurs SNPs diagnostiques pour développer des marqueurs KASPar utilisés chez une collection de variétés du groupe des agrumes vrais permettra une résolution très fine de ce groupe horticole par la discrimination entre ses différentes espèces.

Les marqueurs SNPs générés par la méthode GBS pourraient avoir de nombreuses utilités. Elargir l'étude pour des analyses de liaison et de détection de gènes jouant un rôle dans l'expression des QTL serait intéressant. La sélection des variétés avec des allèles à intérêt agronomique et résistants à des stress biotiques et abiotiques pourraient être développés et utiliser, ainsi, dans les programmes d'amélioration.

La grande diversité des profils phylogéniques révélés dans notre étude sous-tends probablement une part importante de la diversité phénotypique des groupes horticoles étudiés et ouvre la voie à une approche phylogénomique de la génétique d'association.

Références Bibliographiques

- Abkenar, A.A. Isshiki, S. and Tashino, Y. 2004a. Phylogenetic relationships in the true Citrus fruit trees revealed by PCRRFLP analysis of cp DNA. *Sci. Hortic.* 102:233-242.
- Abkenar, A. Isshiki, S. and Tashiro, Y. 2004b. Maternal inheritance of chloroplast DNA in intergenic sexual hybrids of "true Citrus fruit trees" revealed by PCR-RFLP analysis. *J. Hort. Sci. Biotechnol.* 79:360-363.
- Abkenar, A.A. Mardi, M. and Matsumoto, R. 2011. Comparative analysis of organelle DNAs in citrus to clarify the origin of yuzu (citrus junos sieb. Ex tanaka) and its relationship with other acid citrus. *Acta. Hortic.* 892:6.
- Ahmad, R. Struss, D. and Southwick, S.M. 2003. Development and characterization of microsatellite markers in Citrus. *J. Am. Soc. Hort. Sci.* 128: 584-590.
- Aleza, P. Juárez, J. Ollitrault, P. and Navarro, L. 2009. Production of tetraploid plants of non apomictic citrus genotypes. *Plant. Cell. Reports.* 28.
- Aleza, P. Cuenca, J. Juárez, J. Pina, J.A. and Navarro, L. 2010. 'Garbí' Mandarin: a new late-maturing triploid hybrid. *Hort. Science.* 45: 139-141.
- Aleza, P. Cuenca, J. Hernández, M. Juárez, J. Navarro L. and Ollitrault, P. 2015. Genetic mapping of centromeres in the nine Citrus clementinachromosomes using half-tetrad analysis and recombination patterns in unreduced and haploid gametes. *BMC. Plant. Biol.* 15:80
- Alonso, V. and Howell, E. 1983. Comparative-study of the hybrids of the citrus species, tangor temple and ortanique. *Fruits.* 38: 647-652.
- Al-Sadi, A.M. Al-Moqbali, H.S. Al-Yahyai, R.A. and Al-Said, F.A. 2012. AFLP data suggest a potential role for the low genetic diversity of acid lime (Citrus aurantifolia Swingle) in Oman in the outbreak of witches' broom disease of lime. *Euphytica.* 188: 285-297.
- Altschul, S.F. Thomas, L.M. Alejandro, A.S. Jinghui, Z. Zheng, Z. Webb, M. and David, J.L. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic. Acid. Res.* 25: 3389-3402.
- Amar, M.H. Biswas, M.K. Zhang, Z.W. and Guo, W.W. 201. Exploitation of SSR, SRAP and CAPS-SNP markers for genetic diversity of Citrus germplasm collection. *Scientia Horticulturae.* 128: 220-227.

- Araujo, E.F. Queiroz, L.P. and Machado, M.A. 2003. Taxonomic implications from a study of cp-DNA evolution in the tribe Citreae Rutaceae subfamily Aurantioideae. *Org Divers Evol.* 3: 55-62.
- Arumuganathan, K. and Earle, D.E. 1991. Nuclear DNA content of some important plant species. *Plant. Mol. Biol. Rep.* 9(3): 208-218
- Asíns, M.J. Juarez, J. Pina, J.A. Puchades, J. Carbonel, E.A. and Navarro, L. 2002. Una nueva clementina de baja fertilidad llamada Nulesín. *Vida. Rural.* 157: 50-52.
- Asins, M.J. Bernet, G.P. Ruiz, C. Cambra, M. Guerri, J. and Carbonell, E.A. 2004. QTL analysis of Citrus tristeza virus–citradia interaction. *Theor. Appl. Genet.* 108: 603–611.
- Asins, M.J. Fernández, J. Guillermo, R. Bernet, P. Gadea, J. Cambra, et al. 2011. The position of the major QTL for Citrus tristeza virus resistance is conserved among Citrus grandis, C. aurantium and Poncirus trifoliata. *Molecular breeding.* Volume 29, pp 575-587
- Aubert, B. 2001 Genèse et développement de la culture des agrumes et patrimoine génétique méditerranéens de l'histoire naturelle des orangers. Dans la réédition de l'Histoire Naturelle des Orangers- Risseau et Poiteau, tome 2 Connaissance et mémoires éditeur, Paris.
- Baraket, G. Saddoud, O. Chatti, K. Mars, M. Marrakchi, M. Trifi, M. and Salhi Hannachi, A. 2009. Chloroplast DNA analysis in Tunisian Fig cultivars Ficus carica L.: sequence variations of the trnL-trnF intergenic spacer. *Biochem. Syst. Ecol.* 36: 828-835.
- Barboni, T. (2006). Contribution de méthode de la chimie analytique à l'amélioration de la qualité de fruit et à la détermination de mécanismes (EGE) et de risque d'incendie. Thèse, Université de Corse, pp 38-40
- Barkley, N.A. Roose, M.L. Krueger, R.R. and Federici, C.T. 2006. Assessing genetic diversity and population structure in a citrus germplasm collection utilizing simple sequence repeat markers SSRs. *Theor. Appl. Genet.* 112: 1519-1531.
- Barrett, H.C. and Rhodes, A.M. 1976. A numerical taxonomic study of affinity relationships in cultivated Citrus and its close relatives. *Syst. Bot.* 1: 105-136.
- Bastianel, M. Schwarz, S.F. Coleta Filho, H.D. Lin, L.L. Marcos, M. and Koller, O.C. 1998. Identification of zygotic and nucellar tangerine seedlings (Citrus spp.) using RAPD. *Genet. Mol. Biol.* 21: 123-127.

- Bausher, M. Singh, N. Lee, S. Jansen, R. and Daniell, H. 2006. The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var 'Ridge Pineapple': organization and phylogenetic relationships to other angiosperms. *BMC. Plant. Biol.* 6: 21.
- Bayer, R.j. Mabberley, D.j. Morton, C.M. Cathy, H. Sharma, I.K. and Pfeil, B.E. et al. 2009. A molecular phylogeny of the orange subfamily Rutaceae: Aurantioideae using nine cpDNA sequences. *Am. J. Bot.* 96: 668-685.
- Beissinger, T.M. Hirsch, C.N. Sekhon, R.S. Foerster, J.M. Johnson, J.M. Muttoni, G. et al. 2013. Marker density and read depth for genotyping populations using genotyping-by-sequencing. *Genetics*. 193: 1073–1081. doi: 10.1534/genetics.112.147710.
- Ben Yahmed, J. Novillo, P. Garcia-Lor, A. Salvador, A. Ben Mimoun, M. Luro, F. Talon, M. Ollitrault, P. and Morillon, R. 2015. Salt tolerance traits revealed in mandarins (*Citrus reticulata* Blanco) are mainly related to root-to-shoot Cl translocation limitation and leaf detoxification processes. *Scientia Horticulturae*. 191: 90-100.
- Bernardi, J. Licciardello, C. Patrizia Russo, M. et al. 2010. Use of a custom array to study differentially expressed genes during blood orange (*Citrus sinensis* L. Osbeck) ripening. *J. Plant. Physiol.* 167: 301-310.
- Bernardi, J. Mazza, R. Caruso, P. Reforgiato Recupero, G. Marocco, A. and Licciardello, C. 2013. Molecular Breeding; Use of an expressed sequence tag-based method for single nucleotide polymorphism identification and discrimination of *Citrus* species and cultivars. *Mol. Breed.* 31: 705-718.
- Birky, C.W. 2001. The inheritance of genes in mitochondria and chloroplasts, mechanisms and models. *Annu. Rev. Genet.* 35: 125-158.
- Biswas, M.K. Xu, Q. and Deng, X. 2010. Utility of RAPD, ISSR, IRAP and REMAP markers for the genetic analysis of *Citrus* spp. *Sci. Hort.* 124: 254-261.
- Biswas, M.K. Chai, L. Amar, M.H. Zhang, X. and Deng, X. 2011. Comparative analysis of genetic diversity in *Citrus* germplasm collection using AFLP, SSAP, SAMPL and SSR markers. *Sci. Hort.* 129: 798-803.
- Biswas, M.K. Chai, L. Mayer, C. Xu, Q. Guo, W. and Deng, X. 2012. Exploiting BAC-end sequences for the mining, characterization and utility of new short sequences repeat (SSR) markers in *Citrus*. *Mol. Biol. Rep.* 39: 5373-5386.

- Botstein, D. and Risch, N. 2003. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat. Genet.* 33: 228-237.
- Bretó, M.P. Ruiz, C. Pina, J.A. and Asins, and M.J. 2001. The diversification of Citrus clementina Hort. ex Tan., a vegetatively propagated crop species. *Mol phylogenet Evol.* 21: 285-293.
- Brookes, A.J. 1999. The essence of SNPs. *Gene.* 234: 177-186.
- Carbonell-Caballero, J. Alonso, R. Ibanez, V. Terol, J. Talon, M. and Dopazo, J. 2015. A phylogenetic analysis of 35 chloroplast genomes elucidates the relationships between wild and domestic species within the genus Citrus. *Mol. Biol. Evol.* 32: 2015-2035.
- Caruso, M. Federici, C.T. Roose, and M.L. 2008. EST-SSR markers for asparagus genetic diversity evaluation and cultivar identification. *Mol. Breed.* 21: 195-204.
- Castle, W.S., 1987. Citrus rootstock, R.C. Romand R.F. Carlson (Eds.), Rootstocks for fruit crops. Wiley, New York, pp. 361-399.
- Chai, L. Biswas, M.K. Yi, H. Guo, W. and Deng, X. 2013. Transferability, polymorphism and effectiveness for genetic mapping of the Pummelo (Citrus grandis Osbeck) EST-SSR markers. *Sci. Hort.* 155: 85-91.
- Chen, C. Zhou, P. Choi, Y.A. Huang, S. and Gmitter, F.G. 2006. Mining and characterizing microsatellites from citrus ESTs. *Theor. Appl. Genet.* 112: 1248-1257.
- Chen, Z.J. 2007. Genetic and Epigenetic Mechanisms for Gene Expression and Phenotypic Variation in Plant Polyploids. *Annu. Rev. Plant. Biol.* 58: 377-406.
- Chen, C.X. Bowman, K.D. and Choi, Y.A. et al. 2008. EST-SSR genetic maps for Citrus sinensis and Poncirus trifoliata. *Tree. Genet. Genomes.* 4: 1-10.
- Chen, C. and Gmitter, F.G. 2013. Mining of haplotype-based expressed sequence tag single nucleotide polymorphisms in citrus. *BMC. Genomics.* 14: 746.
- Cheng, F.S. and Roose, M.L. 1995. Origin and inheritance of dwarfing by the Citrus rootstock Poncirus trifoliata 'Flying Dragon'. *J. Am. Soc. Hort. Sci.* 120: 286-291.
- Cheng, Y. Vincente, M.C. Meng, H. Guo, W. and Tao, N. 2005. A set of primers for analyzing chloroplast DNA diversity in Citrus and related genera. *Tree. Physiol.* 25: 661-672.

- Clegg, M.T. Gaut, B.S. Learn, G.H. and Morton. B.R. 1994. Rates and patterns of chloroplast DNA evolution. *Proc. Natl. Acad. Sci. USA* 91:6795–6801.
- Coile, N.C. (1995). Notes on nomenclature of Citrus and some related genera, Part 1: Nomenclature of Citrus and other genera in Subfamily Aurantoideae (Rutaceae). *Botany contribution*. 33:4-5
- Collins, F.S. Guyer, M.S. and Charkravarti, A. 1997. Variations on a theme: cataloging human DNA sequence variation. *Science*. 278:1580-1581.
- Combrink, N.K. Labuschagne, M.T. and Bijzet, Z. 2013. Variation of fruit size and shape in Kiyomi tangor families. *Sci. Hort.* 162: 357-364.
- Corazza-Nunes, M. Machado, M. Nunes, W. Cristofani, M. and Targon, M. 2002. Assessment of genetic variability in grapefruits (*Citrus paradisi* Macf.) and pummelos (*C. maxima* (Burm.) Merr.) using RAPD and SSR markers. *Euphytica*. 126: 169-176.
- Cristofani-Yaly, M. Novelli, V.M. Bastianel, M. and Machado, M.A. 2011. Transferability and level of heterozygosity of microsatellite markers in Citrus species. *Plant. Mol. Biol. Rep.* 29: 418-423.
- CTA.com.tn <http://www.cta.com.tn/>
- Cuenca, J. Aleza, P. Navarro, L. and Ollitrault, P. 2013a. Assignment of SNP allelic configuration in polyploids using competitive allele-specific PCR: application to citrus triploid progeny. *Ann. Bot-London*. 111: 731-742.
- Cuenca, J. Aleza, P. Vicent, A. Brunel, D. Ollitrault, P. and Navarro, L. 2013b. Genetically based location from triploid populations and gene ontology of a 3.3-mb genome region linked to alternaria brown spot resistance in citrus reveal clusters of resistance genes. *PloS one* 8: e76755.
- Cuppen, E. Genotyping by Allele-Specific Amplification (KASPar). *CSH Protoc.* 2007:pdb.prot4841.
- Curk, F. 2014. Citrus gene pool organization and nuclear genomic interspecific admixture of cultivated citrus. Corsica, France.
- Curk, F. Ancillo, G. Garcia-Lor, A. Luro, F. Perrier, X. and Jacquemoud-Collet, J.P. et al. Next generation haplotyping to decipher nuclear genomic interspecific admixture in Citrus species: analysis of chromosome 2. *BMC. Genet.* 2014:15: 152.

- Curk, F. Ancillo, G. Ollitrault, F. Perrier, X. Jacquemoud-Collet, J.P. Garcia-Lor, A. et al. 2015. Nuclear species-diagnostic snp markers mined from 454 amplicon sequencing reveal admixture genomic structure of modern citrus varieties. *Plos ONE*. 10: 5: e0125628.
- Curk, F. Ollitrault, F. Garcia-Lor, A. Luro, F. Navarro, L. and Ollitrault, P. 2016. Phylogenetic origin of limes and lemons revealed by cytoplasmic and nuclear markers. *Ann Bot-London*.
- Darlu P. and Tassy P. (1993) La reconstruction phylogénétique. Paris: Masson
- Davey, J. and Blaxter, M.L. 2011. RADSeq: next-generation population genetics. *Brief. Funct. Genomics*. 10:2:108.
- Deng, Z.N. Gentile, A. Nicolosi, E. Continella, G. and Tribulato, E. 1996. Parentage determination of some citrus hybrids by molecular markers. *Proc. Int. Soc. Citricul*. 2: 849-854.
- Deng, Z. La Malfa, S. Xie, Y. Xiong, X. and Gentile, A. 2007. Identification and evolution of chloroplast uni- and trinucleotide sequence repeats in citrus. *Sci. Hortic-Amsterdam*. 111: 186-192.
- Deschamps, S. Llaca, V. and May, G.D. 2012. Genotyping-by-Sequencing in Plants. *Biology* 1:3:460- 83
- Distefano, G. Caruso, M. Casas, G.L. et al. 2011. Maturation in Fruits of Clementine: Biochemical and Molecular Analysis. *Acta. Hortic*. 892: 329-334.
- Distefano, G. La Malfa, S. Gentile, A. and Wu, S. 2013. EST-SNP genotyping of citrus species using high-resolution melting curve analysis. *Tree. Genet. Genomes*. 9: 1271-1281.
- Dong, J. GuangYan, Z. and Hong Q. 2006. Analysis of microsatellites in citrus unigenes. *Acta. Genet. Sinica*. 33: 345-353.
- Dong, J. Qing-liang, Y. Fu-sheng, W. and Li, C. 2010. The mining of citrus EST-SNP and its application in cultivar discrimination. *Agric. Sci. China*. 9: 179-190.
- Duran, C. Appleby, N. Vardy, M. Imelfort, M. Edwards, D. and Batpey, J. 2009. Single nucleotide polymorphism discovery in barley using autoSNPdb. *Plant. Biotech. J*. 7:326-333.
- Durham, R.E. Liou, P.C. Gmitter, F.G. J.R. and Moore, G.A. 1992. Linkage of restriction fragment length polymorphisms and isozymes in Citrus. *Theor. Appl. Genet*. 84: 39-48.
- Elleuch, A. Marrakchi, M. Fakhfakh, H. Levesque, D. Bessais, N. Perreault, J.P. 2003.

- El-Mouei, R. Choumane, W. Dway, F. 2011. Characterization and Estimation of Genetic Diversity in Citrus Rootstocks. *Int. J. Agri. Biol.* 13.
- Elshire, R.J. Glaubitz, Q. Sun, J. Poland, K. Kawamoto, E. Buckler, and Mitchell, S. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species, *PLOS ONE*. 6: e19379
- Engler A. Rutaceae. In: Engler A, Prantl K editors. *Die Natürlichen Pflanzenfamilien?*. Germany: Engelmann Leipzig; 1931. pp. 187-359.
- Fanciullino, A.L. Dhuique-Mayer, C. Luro, F. Casanova, J. Morillon, R. and Ollitrault, P. 2006. Carotenoid diversity in cultivated citrus is highly influenced by genetic factors. *J. Agr. Food Chem.* 54: 4397-4406.
- Fanciullino, A.L. Dhuique-Mayer, C. Luro, F. Morillon, R. Ollitrault, P. 2007. Carotenoid biosynthetic pathway in the citrus genus: number of copies and phylogenetic diversity of seven genes. *J. Agr. Food. Chem.* 55: 7405-7417.
- Fang, D.Q. Zhang, W.C. and Xiao, S.Y. 1993. Studies on taxonomy and evolution of Citrus and its related genera by isozyme analysis in Chinese with English abstract. *Acta. Phytotaxon. Sin.* 31: 329-352.
- Fang, D.Q. Roose, M.L. Krueger, R.R. and Federici, C.T. 1997. Fingerprinting trifoliate orange germ plasm accessions with isozymes, RFLPs, and inter-simple sequence repeat markers. *Theor. Appl. Genetics.* 95: 211-219.
- Fang, D.Q. Krueger, R.R. Roose, M.L. 1998. Phylogenetic relationships among selected Citrus germplasm accessions revealed by inter-simple sequence repeat (ISSR) markers. *J. Am. Soc. Hortic. Sci.* 123: 612-617.
- Fang, J. Song, C. Qian, J. et al. 2010. Variation of cytosine methylation in 57 sweet orange cultivars. *Acta Physiol. Plantarum.* 32: 1023-1030.
- FAO. 2014. The International Treaty on Plant Genetic Resources for Food and Agriculture.
- Federici, C.T. Fang, D.Q. Scora, R.W. and Roose, M.L. 1998. Phylogenetic relationships within the genus Citrus (Rutaceae) and related genera as revealed by RFLP and RAPD analysis. *Theor. Appl. Genet.* 96: 812-822.
- Feng, J. Zhao, X. Laroche, A. Zhen-Xiang, L. Kui, H. and Ziqin, Li. 2014. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Frontiers in plant sciences.* doi: 10.3389/fpls.2014.00484

- Forment, J. Gadea, J. Huerta, L. et al. 2005. Development of a citrus genome-wide EST collection and cDNA microarray as resources for genomic studies. *Plant. Mol. Biol.* 57: 375-391.
- Froelicher, Y. Dambier, D. Bassene, J.B. Costantino, G. Lotfy, S. Didout, C. Beaumont, V. Brottier, P. Risterucci, A.M. Luro, F. and Ollitrault, P. 2008. Characterization of microsatellite markers in mandarin orange (*Citrus reticulata* Blanco). *Mol. Ecol. Res.* 8(1): 119- 122.
- Froelicher, Y. Mouhaya, W. Bassene, J.B. et al. 2011. New universal mitochondrial PCR markers reveal new information on maternal citrus phylogeny. *Tree. Genet. Genomes.* 7: 49-61.
- Fu, Y.X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking, and background selection. *Genetics.* 147: 915-925.
- Fu, Y.X. and Li, W.H. 1993. Statistical tests of neutrality of mutations. *Genetics.* 133, 693-709.
- Fujii, H. Shimada, T. Sugiyama, A. Nishikawa, F. Endo, T. Nakano, M. Ikoma, Y. Shimizu, T. and Omura, M. 2007 Profiling ethylene-responsive genes in mature mandarin fruit using a citrus 22K oligoarray. *Plant. Science.* 173: 340–348.
- Fujii, H. Shimada, T. Sugiyama, A. Endo, T. Nishikawa, F. Nakano, M. Ikoma, Y. Shimizu, T. and Omura, M. 2008 Profiling gibberellin (GA3)-responsive genes in mature mandarin fruit using a citrus 22K oligoarray. *Sci. Hortic.* 116: 291-298
- Fujii, H. Shimada, T. Nonaka, K. Kita, M. Kuniga, T. Endo, T. Ikoma, Y. and Omura, M. 2013. Highthroughput genotyping in citrus accessions using an SNP genotyping array. *Tree. Genet. Genomes.* 9:1:145-153.
- Garcia-Lor, A. Luro, F. Navarro, L. and Ollitrault, P. 2012. Comparative use of Indel and SSR markers in deciphering the interspecific structure of cultivated Citrus genetic diversity: a perspective for genetic association studies. *Mol. Genet. Genomics.* 287: 77-94.
- Garcia-Lor, A. Curk, F. Snoussi-Trifa, H. Morillon, R. Ancillo, G. Luro, F. et al. 2013. A nuclear phylogenetic analysis: SNPs, indels and SSRs deliver new insights into the relationships in the 'true citrus fruit trees' group (Citrinae, Rutaceae) and the origin of cultivated species. *Ann. Bot-London.* 111: 1-19.

- Garcia-Lor, A. Luro, F. Ollitrault, P. and Navarro, L. 2015. Genetic diversity and population structure analysis of mandarin germplasm by nuclear, chloroplastic and mitochondrial markers. *Tree. Genet. Genomes*. 11: 123.
- Garg, K. Green, P. and Nickerson, D.A. 1999. Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags. *Genome. Res*. 9: 1087-1092.
- Ghorabaie, H.R. Ghazvini, R.F. Golein, B. and Nabipour, A.R. 2010. Identification of some Citrus accessions in a citrus germplasm utilizing simple sequence repeat markers (SSRs). *Hortic. Env. Biotech*. 51: 343-347.
- GIFruits.com le secteur des agrumes en Tunisie http://gifruits.com/?page_id=1530&lang=fr
- Glaubitz, J.C. Casstevens, T.M. Lu, F. Harriman, J. Elshire, R.J. et al. 2014 TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *PLoS ONE* 9(2): e90346. doi: 10.1371/journal.pone.0090346
- Gmitter, F.G. Grosser, J.W. Moore, G.A. 1992. Citrus. In: F Hammerschlag, RL Litz, eds. *Biotechnology of perennial crops*. CAB Intl.: Wallingford, Oxon, U.K., 335-369.
- Gmitter, F.G. Xiao, S.Y. Huang, S. Hu, X.L. Garnsey, S.M. and Deng, Z. 1996. A localized linkage map of the citrus tristeza virus resistance gene region. *Theor. Appl. Genet*. 92.
- Gmitter, F.G. Chen, C. Machado, M.A. de Souza, A.A. Ollitrault, P, Froehlicher, Y. and Shimizu, T. 2012. Citrus genomics. *Tree. Genet. Genomes*. 8(3): 611-626.
- Golein, B. Talaie, A. Zamani, Z. Ebadi, A. and Behjatnia, A. 2005. Assessment of genetic variability in some Iranian sweet oranges (*Citrus sinensis* [L.] Osbeck) and mandarins (*Citrus reticulata* Blanco) using SSR markers. *Int. J. Agric. Biol*. 7: 167-170.
- Golein, B. Bigonah, M. Azadvar, M. Golmohammadi, M. 2012. Analysis of genetic relationship between 'Bakraee'(Citrus sp.) and some known Citrus genotypes through SSR and PCR-RFLP markers. *Sci. Hortic*. 148: 147-153.
- Green, R.M. Vardi, A. and Galun, E. 1986. The plastome of Citrus. Physical map, variation among Citrus cultivars and species and comparison with related genera. *TAG.Theor. Appl. Genetics*. 72: 170-177.
- Gulsen, O. and Roose, M.L. 2001a. Determination of genetic diversity and phylogenetic relations to citrus ancestors in lemons by DNA markers. *Bahce*. 30: 53-63.

- Gulsen, O. and Roose, M.L. 2001b. Chloroplast and nuclear genome analysis of the parentage of lemons.. 126: 210-215.
- Gulsen, O. and Roose, M.L. 2001c. Lemons: diversity and relationships with selected Citrus genotypes as measured with nuclear genome markers. *J. Am. Soc. Hortic. Sci.* 126: 309-317.
- Gulsen, O. Uzun, A. Canan, I. Seday, U. and Canihos, E. 2010. A new citrus linkage map based on SRAP, SSR, ISSR, POGP, RGA and RAPD markers. *Euphytica*. 173: 265-277.
- Hadrich, B. Dahak, K. Abdenouri, N. and Kechaou, N. 2008. Etude de séchage des feuilles de bigaradier Revue des Energies Renouvelables SMSTS'08 Alger. 145 – 146
- Hall, T.A. 1999.B ioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic. Acids. Symp.* 41: 95-98.
- Hamdi, I. Elleuch, A. Bessaies, N. and Fakhfakh, H. 2011. Insights on genetic diversity and phylogenetic analysis of Hop stunt viroid (HSVd) population from symptomatic citrus tree in Tunisia. *African Journal of Microbiology Research*. Vol 5(21), pp: 3422-3431.
- Handa, T. Ishizawa, Y. and Oogaki, C. 1986. Phylogenic study of Fraction I protein in the genus Citrus and its close related genera. *Japan. J. Genet.* 61: 15-24.
- Hang, L. Xiaoming, Y. Lianshu, Z. Hualin, Y. Lijun, C. and Xiuxin, D. 2015. Parentage analysis of natural Citrus hybrid 'Zhelong Zhoupigan' based on nuclear and chloroplast SSR markers. *Sci. Hortic-Amsterdam*. 186: 24-30.
- Harley, I.M. Richard, S.B. Virginia, E.S. Debora, W. and Craig, R.E. (2006). Citrus and Fortunella (Kumquat) Rutaceae (rue family). Species profiles Island Agroforestry. Ver 2.1
- Herrero, R. Asins, M.J. Pina, J.A. Carbonell, E.A. and Navarro, L. 1996. Genetic diversity in the orange subfamily Aurantioideae. II. Genetic relationships among genera and species. *Theor. Appl. Genet.* 93: 1327-1334.
- Herrero, R. Asins, M.J. Pina, J.A. Carbonell, E.A. and Navarro, L. 1997 Genetic diversity in the orange subfamily Aurantioideae. II. Genetic relationships among genera and species. *Theor. Appl. Genet.* 93: 1327–1334.
- Hoskins, R.A. Phan, A.C. Naeemuddin, M. et al. 2001. Single nucleotide polymorphism markers for genetic mapping in *Drosophila melanogaster*. *Genome. Research*. 11: 1100-1113.

- Hou, X. Liu, S. Khan, M. Hu, C. and Zhang, J. 2014. Genome-Wide Identification, Classification, Expression Profiling, and SSR Marker Development of the MADS-Box Gene Family in Citrus. *Plant. Mol. Biol. Rep.* 32: 28-41.
- Hudson, R.R. Salatkina, M. and Maddison, W.P., 1992. Estimation of levels of gene flow from DNA sequence data. *Genetics*. 132 :2: 583-590.
- Hynniewta, M. Kumar, S. Malik, S. and Rama, R. 2014. Genetic diversity and phylogenetic analysis of Citrus L. from north-east India as revealed by meiosis, and molecular analysis of internal transcribed spacer region of Rdna. *Meta. Gene*. 2: 237-251.
- Iwamasa, M. Ueno, I. and Nishiura, M. 1967. Inheritance of nucellar embryony in Citrus. *Bull. Fruit. Tree. Res. Stn. E*. 2:9-24.
- Jacquemond, C. Curk, F. Froelicher, Y. and Luro, F. 2013a. Variétés et porte-greffes : création, description et sélection. p. 37-107. In: C. Jacquemond F. Curk and M. Heuzet (eds.), Les clémentiniers et autres petits agrumes, Quae, Versailles.
- Jacquemond, C. Curk, F. and Heuzet, M. 2013b. Créer des plants : pépinière et surgreffage. In: C Jacquemond, F Curk, M Heuzet, eds. Les clémentiniers et autres petits agrumes. Quae: Versailles, 183-207.
- Jaeger, B. Goldbach, H. and Sommer, K. 2000. Release from lime induced iron chlorosis by CULTAN in fruit trees and its characterisation by analysis. *Acta. Hortic*. 531: 107-113.
- Jarrell, D.C. Roose, M.L. Traugh, S.N. and Kupper, R.S. 1992. A genetic map of citrus based on the segregation of isozymes and RFLPs in an intergeneric cross. *Theor. Appl. Genet.* 84: 49-56.
- Jena, S.N. Kumar, S. Nair, K.N. 2009. Molecular phylogeny in Indian Citrus L. Rutaceae inferred through PCR-RFLP and trnL-trnF sequence data of chloroplast DNA. *Sci. Hortic-Amsterdam*. 199: 403-416.
- Jukes T.H., Cantor C.R., 1969. Evolution of protein molecules. In Munro, H.N (Eds.), Mammalian Protein Metabolism. Academy Press., New York, pp. 31-132.
- Jung, Y. Kwon, H. Kang, S. Kang, J. Kim S. 2005. Investigation of the phylogenetic relationships within the genus Citrus (Rutaceae) and related species in Korea using plastid trnL-trnF sequences. *Sci. Hortic*. 104: 179-188.

- Kacar, Y.A. Simsek, O. Mendi, Y.Y. et al. 2012. Investigation of the Self Incompatibility Mechanism in Clementine (*Citrus clementina* Hort. ex Tan.) Using cDNA-AFLP. *Acta. Hortic.* 929: 233-236.
- Kacar, Y. Uzun, A. Polat, I. et al. 2013. Molecular characterization and genetic diversity analysis of mandarin genotypes by SSR and SRAP markers. *Int. J. Food. Agric. Env.* 11: 516-521.
- Kaçar, Y.A. Imek, Ö. Dönmez, D. Boncuk, M. Yeilolu, T. and Ollitrault, P. 2014. Genetic relationships of some Citrus genotypes based on the candidate iron chlorosis genes. *Turk. J. Agric. For.* 38: 340-347.
- Kamiri, M. Stift, M. Srairi, I. et al. 2011. Evidence for non-disomic inheritance in a Citrus interspecific tetraploid somatic hybrid between *C. reticulata* and *C. limon* using SSR markers and cytogenetic analysis. *Plant Cell Reports*.
- Kanno, A. and Hirtai, A. 1993. A transcription map of the chloroplast genome from rice *Oryza sativa*. *Curr. Genet.* 23: 166-174.
- Kepiro, J. (2004). Molecular genetic analyses of nucellar embryony (apomixis) in *Citrus maxima* x *Poncirus trifoliata*. pp. 220. Ph.D. Dissetation, University of California, Riverside
- Kepiro, J.L. and Roose, M.L. 2009. AFLP markers closely linked to a major gene essential for nucellar embryony (apomixis) in *Citrus maxima* x *Poncirus trifoliata*. *Tree. Genet. Genomes.* 6: 1-11.
- Kepiro, J.L. and Roose, M.L. 2010. AFLP markers closely linked to a major gene essential for nucellar embryony (apomixis) in *Citrus maxima* x *Poncirus trifoliata*. *Tree. Genet. Genomes.* 6: 1-11.
- Khalaf AA, Jones J, Moore GA and Gmitter FG. Microarray expression analysis in Nagami Kumquat in response to canker infection, in Plant and Animal Genome XV conference, p.485, San Diego, Calif, USA, January 2007
- Khouranga, M. Babaeia, A. Sefidkonb, F. Naghavic, M.R. Asgaric, D. and Potterd, D. 2014. Phylogenetic relationship in *Fritillaria* spp. of Iran inferred from ribosomal ITS and chloroplast trnL-trnF sequence data. *Bioch. Syst. Ecol.* 57: 451-457.
- Kijas, J. Fowler, J. and Thomas, M. 1995. An evaluation of sequence tagged microsatellite site markers for genetic analysis within Citrus and related species. *Genome.* 38: 349-355.

- Kijas, J.M.H. Thomas, M.R. Fowler, J.C.S. and Roose, M.L. 1997. Integration of trinucleotide microsatellites into a linkage map of Citrus. *Theor. Appl. Genet.* 94: 701-706.
- Kilpinen H, Barrett JC (2013) How next-generation sequencing is transforming complex disease genetics. *Trends Genet* 29: 23-30. Available: <http://www.ncbi.nlm.nih.gov/pubmed/23103023>. Accessed 11 November 2013.
- Kimura, M. 1983. The Neutral Theory of Molecular Evolution.
- Kimura, M. and Crow, J. 1964. The number of alleles that can be maintained in a finite population. *Genetics*. 49: 725-738.
- Kobayashi, S. Ikeda, I. and Nakatani, M. 1967a. Studies on the nucellar embryogenesis in Citrus. I. Formation of nucellar embryo and development of ovule. *Bull. Fruit. Tree. Res. Stn. E* 2:9-24
- Kobayashi, S. Ikeda, I. and Nakatani, M. 1967b. Studies on the nucellar embryogenesis in Citrus. II. Formation of the promordium cell of the nucellar embryo in the ovule of the flower bud, and its meristematic activity. *J. Jpn. Soc. Hortic. Sci.* 48:179-185.
- Kohany, O. Gentles, A.J. Hankus, L. and Jurka, J. 2006. Annotation, submission and screening of repetitive elements in Repbase: Rep base Submitter and censor. *BMC. Bioinf.* 7, 474.
- Krueger, R.R. and Navarro, L. Citrus germplasm resources. In: Khan I, editor. Citrus genetics, breeding and biotechnology. Wallington: CABI Publishing; 2007. pp. 45-140.
- Krug, C. 1943. Chromosome number in the subfamily Aurantioideae with special reference to the genus Citrus. *Bot. Gaz.* 104: 602-611.
- Kumar, S. Nair, K.N. and Jena, S. 2013. Molecular differentiation in Indian Citrus L. (Rutaceae) inferred from nrDNA ITS sequence analysis. *Genet. Resour. Crop. Ev.* 60: 59-75.
- Langmead B1, Salzberg SL. 2012. Nat Methods. Fast gapped-read alignment with Bowtie 2. 9(4):357-9. doi: 10.1038/nmeth.1923.
- Lecointre et al., 2001 : Lecointre G., Le Guyader H. & Visset D. (2001) Classification phylogénétique du vivant. Berlin: Belin.
- Lee, J.T. 2003. Molecular biology: complicity of gene and pseudogene. *Nat.* 42: 26-28.

- Lee, Y.G. Jeong, N. Kim, J.H. Lee, K. Kim, K.H. Pirani, A. et al. 2015. Development, validation and genetic analysis of a large soybean SNP genotyping array. *Plant. J.* 81: 625-636.
- Li, Y.Z. Cheng, Y.J. Yi, H.L. and Deng, X.X. 2006. Genetic diversity in mandarin landraces and wild mandarins from China based on nuclear and chloroplast simple sequence repeat markers. *J. Hort. Sci. Biotechnol.* 81: 371-378.
- Li, Y. Cheng, Y. Tao, N. and Deng, X. 2007. Phylogenetic Analysis of Mandarin Landraces, Wild Mandarins, and Related Species in China Using Nuclear LEAFY Second Intron and Plastid trnL-trnF Sequence. *J. Am. Soc. Hort. Sci.* 1326: 796-806.
- Li, X. Xie, R. Lu, Z. and Zhou, Z. 2010. The Origin of Cultivated Citrus as Inferred from Internal Transcribed Spacer and Chloroplast DNA Sequence and Amplified Fragment Length Polymorphism Fingerprints. *J. Amer. Soc. Hort. Sci.* 1354: 341-350.
- Liang, G. Xiong, G. Guo, Q. He, Q. and Li, X. 2007. AFLP analysis and the taxonomy of Citrus. *Acta. Hort.* 760: 137-142.
- Librado, P. and Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics.* 25: 1451-1452.
- Lijavetzky, D. Cabezas, J.A. Ibanez, A. Rodriguez, V. and Martinez-Zapater, J.M. 2007. High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology. *BMC. Genomics.* 8: 424.
- Ling P, Yu C, Deng Z, et al. 1999. Citrus genome mapping with AFLP markers. Plant & Animal Genomes XIII Conf, San Diego, CA, USA Poster.
- Liu, J. Pang, X. Cheng, Y. Meng, H. and Deng, X. 2002. Molecular characterization of the nuclear and cytoplasmic genomes of intergeneric diploid plants from cell fusion between *Microcitrus papuana* and Rough lemon. *Plant. Cell. Reports.* 21: 327-332.
- Liu, S. Li, W. Long, D. Hu, C. and Zhang, J. 2013. Development and Characterization of Genomic and Expressed SSRs in Citrus by Genome-Wide Analysis. *PLoS ONE.* 8: e75149.
- Lombardoa, G. Schicchia, R. Marinoa, P. and Pallaa, F. 2012. Genetic analysis of Citrus aurantium L. (Rutaceae) cultivars by ISSR molecular markers. *Plant. Biosystems.* 146: 19-26.
- Longley, A.E. Polycary, polyspory and polyploidy in Citrus and Citrus relatives. Journal of the Washington Academy of Science. 1925, 15, 347-351.

- Loffy, S. Luro, F. Carreel, F. Froelicher, Y. Rist, D. and Ollitrault, P. 2003. Application of cleaved amplified polymorphic sequence method for analysis of cytoplasmic genome among Aurantioideae intergeneric somatic hybrids. *J.Am. Soc. Hort. Sci.* 128: 225-230.
- Lu, Z. Zhou, Z. and Xie, R. 2011. Molecular Phylogeny of the "True Citrus Fruit Trees" Group (Aurantioideae, Rutaceae) as Inferred from Chloroplast DNA Sequence. *Agric. Sci. China* 10: 49-57.
- Luro, F. Lorieux, M. Laigret, F. Bove, J.M and Ollitrault, P. 1994. Genetic mapping of an intergeneric Citrus hybrid using molecular markers. *Fruits*. 49: 404-408.
- Luro, F. Laigret, F. Bove, J.M and Ollitrault, P. 1995. DNA amplified fingerprinting, a useful tool for determination of genetic origin and diversity analysis in Citrus. *Hort.Science*. 30: 1063-1067.
- Luro, F. Rist, D. and Ollitrault, P. 2001. Evaluation of genetic relationships in Citrus genus by means of sequence tagged microsatellites. *Acta. Hortic.* 237-242.
- Luro, F. Froelicher, Y. Costantino, G. and Ollitrault, P. 2003. Citrus STMS markers development and mapping. Plant & Animal Genomes XI Conference. Workshop: Citrus,79. Town & Country convention Center, San Diego, CA.
- Luro, F. Costantino, G. Terol, J. et al. 2008. Transferability of the EST-SSRs developed on Nules clementine (Citrus clementina Hort ex Tan) to other Citrus species and their effectiveness for genetic mapping. *BMC. Genomics* 9: 287.
- Luro, F. Gatto, J. Costantino, G. and Pailly, O. 2011. Analysis of genetic diversity in Citrus. *Plant. Genet. Resour.* 9: 218-221.
- Lynch, M. and Crease, T. 1990. The analysis of population survey data on DNA sequence variation. *Mol. Biol. Evol.* 7: 377-394.
- Mabberley, D.J. 1997. A classification for edible Citrus Rutaceae. *Telopea*. 7: 167-172.
- Mabberley, D.J. 2004. Citrus Rutaceae: a Review of recent advances in etymology, systematics and medical applications. *Blumea*. 49: 481-498.
- Mabberley DJ. 2008. Mabberley's plant-book: a portable dictionary of plants, 3rd edn. Cambridge University Press, Avon
- Malik, M.N. Scora, R.W. and Soost, R.K. 1974. Studies on the origin of the lemon. *Hilgardia*. 42 : 361-382.

- Malik, S.K. Uchoi, A. Kumar, S. et al. 2013. Molecular characterization of Citrus macroptera Montr.(Satkara): An endangered wild species from northeast India. *Plant. Biosyst.* 147: 857-863.
- Martinez-Godoy, M.A. Mauri, N. Juarez, J. 2008. A genome-wide 20 K citrus microarray for gene expression analysis. *BMC. Genomics.* 9: 318.
- Maxam, A.M. and Gilber, W. 1977 A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA*, 74: 560-564
- Metzker, M.L. 2010 Sequencing technologies - the next generation. *Nat. Rev. Genet.* 11: 31-46. doi:10.1038/nrg2626. Molecular variability of Citrus exocortis viroid in a single naturally infected citrus tree. *Plant. Protec. Sci.* 1212-2580.
- Mondal B and Saha R. 2014. Identification of zygotic and nucellar seedlings of Citrus reticulata and Citrus aurantifolia using RAPD. *International Journal of Advanced Biotechnology and Research.* Vol5, Issue1, 2014, pp 25-30
- Moore, G.A. Tozlu, I. Weber, C.A. and Guy, C.L. 2000. Mapping quantitative trait loci for salt tolerance and cold tolerance in Citrus grandis (L.) Osb. x Poncirus trifoliata (L.) Raf. hybrid populations. *Acta. Hort.* 535: 37-45.
- Moore, G.A. 2001. Oranges and lemons: clues to the taxonomy of Citrus from molecular markers. *Genetics.* 9: 536-540.
- Morales, M, Roig, E, Monforte, A.J. Arus, P. Garcia-Mas, J. 2004. Single-nucleotide polymorphisms detected in expressed sequence tags of melon (Cucumis melo L.). *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* 47: 352-360.
- Morton, C. Grant, M. and Blackmore, S. 2003. Phylogenetic relationships of the Aurantioideae inferred from chloroplast DNA sequence data. *Am. J. Bot.* 90: 1463-1469.
- Morton. C. 2008. Phylogenetic relationships of the Aurantioideae Rutaceae based on the nuclear ribosomal DNA ITS region and three noncoding chloroplast DNA regions, atpB-rbcL spacer, rps16, and trnL-trnF. *Org. Div. Evol* 9: 52:68.
- Nartvaranant, P. and Nartvaranant, K. 2011. Analysis based on AFLP markers of the genetic variations and their relationships for pummelo cultivars grown in the central region of Thailand. *Songklanakarin Journal of Science and Technology.* 33: 499-508.

- Naz, S. Shahzadi, K. Rashid, S. Saleem, F. Zafarullah, A. and Ahmad, S. 2014. Molecular Characterization and Phylogenetic Relationship of Different Citrus Varieties of Pakistan. *J. Anim. Plant. Sci.* 24: 315-320.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA.* 70: 3321-3323.
- Nei, M. 1977. F-statistics and analysis of gene diversity in sub-divided populations. *Ann. Hum. Genet.* 41: 225 - 233.
- Nei, M. 1987. Molecular Evolutionary Genetics Columbia University Press, New York, NY, USA
- Nei, M. and Tajima, F. 1983. Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. *Genetics.* 105: 207-217.
- Nicolosi, E. Deng, Z.N. Gentile, A. La Malfa, S. Continella, G. and Tribulato, E. 2000. Citrus phylogeny and genetic origin of important species as investigated by molecular markers. *Theor. Appl. Genet.* 100: 1155-1166.
- Novelli, V.M. Cristofani, M. and Machado, M.A. 2000. Evaluation of microsatellite markers in cultivars of sweet orange (*Citrus sinensis* [L.] Osbeck). *Acta. Hortic.* 47-50.
- Novelli, V.M. Takita, M.A. and Machado, M.A. 2004. Identification and analysis of single nucleotide polymorphisms (SNPs) in citrus. *Euphytica.* 138: 227-237.
- Novelli, V.M. Cristofani, M. Souza, A.A. and Machado, M.A. 2006. Development and characterization of polymorphic microsatellite markers for the sweet orange (*Citrus sinensis* L. Osbeck). *Genet. Mol. Biol.* 29: 90-96.
- Ogawa, K. Kawasaki, A. Omura, M. Yoshida, T. Ikomab, Y. and Yano, M. 2001. 3',5'-Di-glucopyranosylphloretin, a flavonoid characteristic of the genus *Fortunella*. *Phytochemistry.* 57: 737-742.
- Olivares-Fuster, O. Hernandez-Garrido, M. Guerri, J. and Navarro, L. 2007. Plant somatic hybrid cytoplasmic DNA characterization by single-strand conformation polymorphism. *Tree. physiol.* 27: 785-792.
- Ollitrault, P. and De Rocca Serra, D. 1992. Les agrumes. In: A Gallais, H Bannerot, eds. Amélioration des espèces végétales cultivées. Inra: Paris, 633-654.

- Ollitrault P, Jacquemond C, Dubois C, Luro F. 2003. Citrus. In: Hamon Perla, Seguin Marc, Perrier Xavier, Glaszmann Jean-Christophe, eds. Genetic diversity of cultivated tropical plants. Cirad: Montpellier, 193-217.
- Ollitrault, F. Terol, J. Pina, J.A. Navarro, L. Talon, M. and Ollitrault, P. 2010. Development of SSR markers from Citrus clementina (Rutaceae) BAC end sequences and interspecific transferability in Citrus. *Am. J. Bot.* 97: e124-9.
- Ollitrault P, Terol J, Chen C, et al. 2011. A reference linkage map of C. clementina based on SNPs, SSRs and Indels. Plant & Animal Genomes XIX Conference P477.
- Ollitrault, P. Terol, J. Garcia-Lor, A. Bérard, A. Chauveau, A. Froelicher, Y. et al. 2012a. SNP mining in C.clementina BAC end sequences; transferability in the Citrus genus (Rutaceae), phylogenetic inferences and perspectives for genetic mapping. *BMC. Genomics*. 13: 13.
- Ollitrault, P. Terol, J. Chen, C. Federici, C.T. Lotfy, S. Hippolyte, I. et al. 2012b. A reference genetic map of C.clementina hort.ex Tan; Citrus evolution inferences from comparative mapping. *BMC. Genomics*. 13: 593.
- Ollitrault P, Navarro L. 2012c. Citrus. In: M Badenes, De Byrne, eds. Fruit Breeding, Dordrecht Heildlberg edn. Springer New York: London, 623-662.
- Ollitrault, F. Terol, J. Martin, A.A. Pina, J.A. Navarro, L. Talon, M. and Ollitrault, P. 2012d. Development of indel markers from Citrus clementina (Rutaceae) BAC-end sequences and interspecific transferability in Citrus. *Am. J. Bot.* 99: e268-73.
- Olmstead, R.G. and Palmer, J.D. 1994. Chloroplast DNA systematics: a review of methods and data analysis. *Am. J. Bot.* 81: 1205-1224.
- Omura M, Ueda T, Kita M, et al. 2000. EST mapping of Citrus. Proceedings of the International Society of Citriculture IX congress. Orlando, USA.
- Pang, X.M. Hu, C.G. and Deng, X.X. 2003. Phylogenetic relationships among Citrus and its relatives as revealed by SSR markers. *Acta. Genetica. Sinica*. 30: 81-87.
- Pang, X. Hu, C.G. and Deng, X.X. 2007. Phylogenetic relationships within Citrus and its related genera as inferred from AFLP markers. *Genet. Resour. Crop. Ev.* 54: 429-436.
- Penjor, T. Anai, T. Nagano, T. Matsumoto, R. and Yamamoto, M. 2010. Phylogenetic relationships of Citrus and its relatives based on rbcL gene sequences. *Tree. Genet. Genomes*. 6: 931-939.

- Penjor, T. Yamamoto, M. Uehara, M. Idel, M. Matsumoto, N. Matsumoto, R. Nagano, Y. 2013. Phylogenetic Relationships of Citrus and Its Relatives Based on MatK Gene Sequences. *Plos. One*. 8(4), e62574.
- Perrier, X. and Jacquemoud-Collet, J.P. 2006. DARwin software. <http://darwin.cirad.fr/darwin>.
- Pessina, D. Gentili, R. Barcaccia, G. 2011. DNA content, morphometric and molecular marker analyses of Citrus limonimeditica, C. limon and C. medica for the determination of their variability and genetic relationships within the genus Citrus. *Sci. Hortic*. 129: 663-673.
- Petit, R.J. Duminil, J. Fineshi, S. Hampe, A. Salvini, D. and Vendramin, G.V. 2005. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol. Ecol*. 14: 689-701.
- Petr, D. Auton, A. Goncalo, A. Cornelis, A. Banks, E. Mark, A. 2011. 1000 Genomes Project Analysis Group. The Variant Call Format and VCFtools. Bioinformatics
- Pfeil, B.E, Crisp, M.D. 2008. The age and biogeography of Citrus and the orange subfamily (Rutaceae: Aurantioideae). *Am. J. Bot*. 95: 1621-1631.
- Poczai, P. Hyvönen, J. 2011. Identification and characterization of plastid trnF(GAA) pseudogenes in four species of Solanum (Solanaceae). *Biotechnol. Lett*. 33: 2317-2323.
- Poczai, P. and Hyvönen, J. 2013. Discovery of novel plastid phenylalanine (trnF) pseudogenes defines a distinctive clade in Solanaceae. *Springer. Plus*. 2, 459.
- Polat, I. Kacar, Y. Yesiloglu, T. et al. 2012. Molecular characterization of sour orange (Citrus aurantium) accessions and their relatives using SSR and SRAP markers. *Genet. Mol. Res*. 11: 3267-3276.
- Pons, O. and Petit, R.J. 1996. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*. 144: 1237-1245
- Quandt, D. Müller, K. Stech, M. Frahm, J.P. Frey, W. Hilu, K.W. Borsch, T. 2004. Molecular evolution of the chloroplast trnL-F region in land plants. *Monogr. Syst. Bot. Missouri. Bot. Gard*. 98: 13-37.
- Ramadugu, C. Pfeil, B.E. Keremane, M.L. Lee, R.F. Maureira-Butler, I.J. and Roose, M.L. 2013. A six nuclear gene phylogeny of Citrus (Rutaceae) taking into account hybridization and lineage sorting. *PloS One*. 8: e68410.

- Rao, M.N Soneji, J.R. Chen, C.X. Huang, S. and Gmitter, F.G. 2008. Characterization of zygotic and nucellar seedlings from sour orange-like citrus rootstock candidates using RAPD and EST-SSR markers. *Tree. Genet. Genomes*. 4: 113-124.
- Reboud, X. and Zeyl, C. 1994 Organelle inheritance in plants. *Heredity*. 72:132–140
- Rogers, A.R. and Harpending, H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9: 552-569.
- Roose, M.L. Feng, D. Cheng, F.S. Tayyar, R.I. Federici, C.T. and Kupper, R.S. 2000. Mapping the Citrus genome. *Acta. Hort.* 25-32.
- Roose, M.L. Isozymes and DNA restriction fragment polymorphism in citrus breeding and systematics. *Proc. Int. Soc. Citricult.* pp. 155-165.
- Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.
- Sakkaa, H. Baraket, G. Dakhlaoui Dkhil, S. Zehdi Azzouzi, S. and Salhi-Hannachi, A. 2013. Chloroplast DNA analysis in Tunisian date-palm cultivars (*Phoenix dactylifera* L.): Sequence variations and molecular evolution of trnL (UAA) intron and trnL (UAA) trnF (GAA) intergenic spacer. *Sci. Hort.* 164: 256-269.
- Samuel, R. Ehrendorfer, F. Chase, M.W. and Greger, H. 2001. Phylogenetic analyses of Aurantoideae (Rutaceae) based on non-coding plastid DNA sequences and phytochemical features. *Plant. Biol.* 3: 77-87.
- Sang, T. Crawford, D.J. and Stuessy, T.F. 1997. Chloroplast DNA phylogeny, reticular evolution and biogeography of *Paeonia* Paeoniaceae. *Am. J. Bot.* 84: 1120-36.
- Sankar, A.A. and Moore, G.A. 2001. Evaluation of inter-simple sequence repeat analysis for mapping in Citrus and extension of the genetic linkage map. *Theor. Appl. Genet.* 102: 206-214.
- Santos, M.G. Passos, O.S. Soares Filho, W.S et al. 2013. Variability analysis of 'Persian' acid lime tree selections using agronomic and molecular markers. *Genet. Mol. Res.* 12: 4604-4614.
- Savita Bhagat, A. Pati, P.K. Virk, G.S. and Nagpal, A. 2012. An efficient micropropagation protocol for Citrus jambhiri Lush. and assessment of clonal fidelity employing anatomical studies and RAPD markers. *In Vitro Cell. Dev-Plant.* 48: 512-520.

- Schroeder, H. and Fladung, M. 2014. Differentiation of *Populus* species by chloroplast SNP markers for barcoding and breeding approaches. *Forest. Biogeosci. Forestry*. 8: 544-546.
- Scora, R.W. 1975. On the history and origin of Citrus. *Bull. Torrey. Bot. Club*. 102: 369-375.
- Scora, R.W. 1988. Biochemistry, taxonomy and evolution of modern cultivated citrus. *Proceedings of the Sixth International Citrus Congress* : 277-289.
- Shimada, T. Hirabayashi, T. Endo, T. Fujii, H. Kita, M. Omura, M. 2005. Isolation and characterization of the somatic embryogenesis receptor-like kinase gene homologue (CitSERK1) from Citrus unshiu Marc. *Sci. Hortic*. 103: 233-238.
- Shimada, T. Fujii, H. Endo, T. et al. 2014. Construction of a citrus framework genetic map anchored by 708 gene-based markers. *Tree. Genet. Genomes*. 1-13.
- Shimizu, T. Yoshioka, T. Nagasaki, H. et al. 2012. Whole genome sequencing and mapping analysis for identifying polymorphism among 11 citrus varieties. XII International Citrus Congress - Valencia, Spain S03O03: 62.
- Simonsen, K.L. Churchill, G.A. and Aquadro, C.F. 1995. Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics*. 141: 413-429.
- Simsek, O. Aka Kacar, Y. Yesiloglu, T and Ollitrault, P. 2011. Determination by SSCP markers of allelic diversity of candidate genes for tolerance to iron chlorosis in citrus germplasm. *Acta. Hortic*. (ISHS) 892: 85-91.
- Singh, A. Sinha, P. Dubey, A. and Srivastav, M. 2011. Characterization of citrus germplasm using simple sequence repeat (SSR) markers. *Indian. J. Genet. Plant. Breed*. 71: 349-355.
- Slatkin, M. and Barton, N.H. 1989. A comparison of three indirect methods for estimating the average level of gene flow. *Evolution*. 43:1349-1368.
- Snoussi Trifa, H., Rim H'lel, B., Bel Falah, Z., Ben Mimoun, M., Jedidi Neji, E., Najjar, A., Duval, M.F., Ollitrault, P. (2006). In : 13 èmes Journée de l'IRESA, 14-15 décembre 2006. S. 1: s.n., 7 p. Journée de l'IRESA. 13, 2006-12-14/2006-12-15, Tunis, Tunisie
- Snoussi, H. Duval, M.F. Garcia-Lor, A. et al. 2012. Assessment of the genetic diversity of the Tunisian citrus rootstock germplasm. *BMC. Genetics*. 13: 16-2156-13-16.
- Sonah, H. Bastien, M. Elmer, I. ATardivel, G. Légaré, B. Boyle, É. Normandeau, J. Laroche, S. Larose, M. Jean, Belzile, F. 2013. An Improved Genotyping by Sequencing (GBS)

Approach Offering Increased Versatility and Efficiency of SNP Discovery and Genotyping. *PLoS One*. 8(1): e54603.

Storey, R. and Walker, R.R. 1999. Citrus and salinity. *Sci. Hortic*. 78: 39-81.

Su, H.J. AHogenhout, S.A. Al-Sadi, A.M, and Kuo, C.H. 2014. Complete chloroplast Genome Sequence of Omani Lime (*Citrus aurantiifolia*) and Comapative Analysis within the Rosids. *Plos one*. 9: e113049.

Swingle, W.T. and Reece P.C. 1967. The botany of Citrus and its wild relatives in the orange subfamily. The Citrus industry. Vol 1. Reuther W, Webber HJ, Batchelor DL (Eds.), Berkeley: University of California, pp 190-340.

Sykes, S.R. and Possingham, J.V. The effect of excluding insect pollinators on seediness of Imperial mandarin fruits. 1992. *Australian. J. Exper. Agric*. 32(3) 409 - 411

Taberlet, P. Gielly, L. Pautou, G. Bouvet, J. 1991. Universal primers for ampli?cation of three non-coding regions of chloroplast DNA. *Plant. Mol. Biol*. 17: 1105-1109.

Tajima, F. 1983 Evolutionary relationship of DNA sequences in finite populations. *Genetics*. 105: 437-460

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 123: 585-595.

Takita, MA. 2015. Exploiting the Citrus genome towards understanding genetic traits related to fruit quality and disease resistance. Plant animal genome XXIII. San Diego CA.

Tamura, K. Stecher, G. Peterson, D. Filipski, A. Kumar, S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol*. 30: 2725-2729.

Tanaka T. 1954. Species problem in Citrus (Revisio Aurantiacearum IX). Japanese Society for Promotion of Science: Tokyo, Japan.

Tanaka T. 1961. Citologia: Semi-centennial Commemoration Papers on Citrus Studies. Citologia Supporting Foundation: Osaka.

Tanaka, T. 1977. Fundamental discussion of Citrus classification. Study. *Citrologia*. 14: 1-6

Tanaka, T. The taxonomy and nomenclature of Rutaceae Aurantioideae. *Blumea*. 1936: 2: 101-110.

- Terol, J. Conesa, A. Colmenero, J.M. Cercos, M. Tadeo, F. Agusti, J. 2007. Analysis of 13000 unique Citrus clusters associated with fruit quality, production and salinity tolerance. *BMC. Genomics*. 8: 31.
- Terol, J. Naranjo, M.A. Ollitrault, P. Talon, M. 2008. Development of genomic resources for Citrus clementina: characterization of three deep-coverage BAC libraries and analysis of 46,000 BAC end sequences. *BMC. Genomics*. 9: 423.
- Terol, J. Carbonell, J. Alonso, R. et al. 2012. Sequencing of 150 citrus varieties: linking genotypes to phenotypes. XII International Citrus Congress - Valencia, Spain S03O02: 61.
- Thorne, R.F. 2000. The classification and geography of the flowering plant: Dicotyledons of the class Angiospermae (subclasses Magnoliidae, Ranunculidae, Caryophyllidae, Dilleniidae, Rosidae, Asteridae, and Lamiidae). *Bot. Rev.* 66: 441-647.
- Torres, A.M. Soost, R.K. and Diedenhofen, U. 1978. Leaf isozymes as genetic markers in Citrus. *Am. J. Bot.* 65: 869-881.
- Tripolitsiotis, C. Nikoloudakis, N. Linos, A. and Hagidimitriou, M. 2013. Molecular Characterization and Analysis of the Greek Citrus Germplasm. *Nat. Bot. Hortic. Agr. Cluj-Napoca*. 41.
- Tshering, P. Anai, T. Nagano, Y. Matsumoto, R. and Yamamoto, M. 2010. Phylogenetic relationships of Citrus and its relatives based on *rbcL* gene sequences. *Tree. Genet. Genomes*. 6: 931-939.
- Tudela, D. and Primo-Millo, E. 1992. 1-Aminocyclopropane-1-Carboxylic Acid Transported from Roots to Shoots Promotes Leaf Abscission in Cleopatra Mandarin (Citrus reshni Hort. ex Tan.) Seedlings Rehydrated after Water Stress. *Plant. Physiol.* 100: 131-137.
- Urasaki, N. Yoshida, K. Uehara, T. Inoue, H. and Onda, S. 2005. Single nucleotide polymorphism in shiikuwasha Citrus depressa Hayata chloroplast DNA, *trnL-trnF*. *Jpn. J. Trop. Agric.* 49: 246-251.
- Uzun, A. Gulsen, O. Seday, U. Yesiloglu, T. Aka-Kacar, Y. and Tuzcu, O. 2011a. Investigation of genetic relationships among trifoliata oranges and their hybrid relatives based on ISSR markers. *Rom. Biotech. Lett.* 16: 64-31.
- Uzun, A. Yesiloglu, T. Polat, I. et al. 2011b. Evaluation of Genetic Diversity in Lemons and Some of Their Relatives Based on SRAP and SSR Markers. *Plant. Mol. Biol. Report.* 29: 693-701.

- Vos, P. Uitdewilligen, J. Voorrips, R. Visser, R. and Van Eck, H. 2015. Development and analysis of a 20K SNP array for potato (*Solanum tuberosum*): an insight into the breeding history. *Theor. Appl. Genet.* 128: 2387-2401.
- Webber, H.J. 1967. History and development of the citrus industry. The citrus industry, Vol 1. Reuther W, Webber HJ, Batchelor LD (Eds.) Berkeley: University of California Press, pp 1-39.
- Wright S (1978) Evolution and the Genetics of Population, Variability Within and Among Natural Populations. The University of Chicago Press, Chicago
- Wright, S. 1965. The interpretation of population structure by F-statistics with special regard to system of mating. *Evolution*. 19.
- Wu, G.A. Prochnik, S. Jenkins, J. et al. 2014. Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nat. biotech.* 32: 656-662.
- Xiao, J. Chen, J. Zhang, H. Xu, H. Wang, H. and Xie, M. 2011. Gene expression profiling in response to drought stress in citrus leaves by cDNA-AFLP. *Acta.* 38: 417-424.
- Xing, C. Schumacher, F.R. Xing, G. Lu, Q. Wang, T. and Elston, R.C. 2005. Comparison of microsatellites, single-nucleotide polymorphisms (SNPs) and composite markers derived from SNPs in linkage analysis. *BMC. Genetics.* 6 :1: S29.
- Xu, Q. Chen, L.L. Ruan, X. Chen, D. Zhu, A. Chen, C. Bertrand, D. 2012. The draft genome of sweet orange (*Citrus sinensis*). *Nat.* 10, 1038.
- Xu, Q. Chen, L. Ruan, X. et al. 2013. The draft genome of sweet orange (*Citrus sinensis*). *Nat. Genet.* 45: 59-66.
- Yakushiji, H. Morinaga, K. and Nonami, H. 1998. Sugar accumulation and partitioning in Satsuma mandarin tree tissues and fruit in response to drought stress. *J. Am. Soc. Hortic. Sci.* 123: 719-726.
- Yamamoto, M. Kobayashi, S. Nakamura, Y. Yamada, Y. 1993. Phylogenic relationships of citrus revealed by RFLP analysis of mitochondrial and chloroplast DNA. *Jpn. J. Breed.* 43: 355-365.
- Yamamoto, M. Kouno, R. Nakagawa, T. Usui, T. and Kubo, T. 2011. Isozyme and DNA analyses of local Citrus germplasm on Amami Islands, Japan. *J. Jpn. Soc. Hortic. Sci.* 80: 268-273.

- Yamamoto, M. Tsuchimochi, Y. Ninomiya, T. Koga, T. Kitajima, A. Yamasaki, A. et al. 2013. Diversity of Chloroplast DNA in Various Mandarins Citrus spp. and Other Citrus Demonstrated by CAPS Analysis. *J. Jpn. Soc. Hortic. Sci.* 2: 106-113.
- Ye, J., Zhu, A. Tao, N. Xu, Q. Xu, J. and Deng, X. 2010. Comprehensive analysis of expressed sequence tags from the pulp of the red mutant 'Cara Cara' navel orange (Citrus sinensis Osbeck). *J. Integ. Plant. Biol.* 52: 856-867.
- Yildiz, E. Kaplankiran, M. Demirkese, T. Uzun, A. and Toplu, C. 2013. Identification of Zygotic and Nucellar Individuals Produced from Several Citrus Crosses Using SSRs Markers. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca.* 41.
- Yingzhi, L. Yunjiang, C. Nengguo, T. and Xiuxin, D. 2007. Phylogenetic analysis of mandarin land races, wild mandarins, and related species in China using nuclear LEAFY second intron and plastid trnL-trnF sequence. *J. Am. Soc. Hort. Sci.* 132: 796-806.
- Zhang, Q. Zhao, H. Lin, L. and Wen, H. 2011. Comparative studies on the quality of different cultivars of Citrus grandis var. tamentosa. *Hort. Med. Plant* 2: 5-8.
- Zhou, C. Qi, Y. You, X. et al. 2013. Leaf cDNA-AFLP analysis of two citrus species differing in manganese tolerance in response to long-term manganese-toxicity. *BMC. Genomics.* 14: 621.
- Zhu, S. Wu, B. Ma, Y. Chen, J. and Zhong, G. 2013. Obtaining citrus hybrids by in vitro culture of embryos from mature seeds and early identification of hybrid seedlings by allele-specific PCR. *Sci. Hortic.* 161: 300-305.

Annexes

Annexe 1 : Liste des séquences des accesions utilisées dans l'analyse *in silico* des SNP

Tribu	Sous-tribu	Genre	Espèce	Clade Bayer <i>et al.</i> (2009)	Numéro des accesions (GenBank)							
					Région codante atpB	Espaceur rbcL-atpB	Espaceur matK-5'trnK	Espaceur psbM-trnDGUC	Intron trnG	Espaceur rps16	Région trnL-F	Espaceur rps4-trnT
Clauseneae	Micromelinae	Micromelum	Micromelum minutum	A,B,E	EF118889	AF320854	EF138904	EF164876	EF176560	AF320266	EF126691	EF134696
	Clauseniaceae	Clausena	Clausena excavata	C,D,E	AF066841	AF320849	EF138881	EF164853	EF176537	AY295258	EF126674	EF134673
			Clausena harmandiana	C,D,E	EF118868	AF320892	EF138882	EF164854	EF176538	EF126608	EF126675	EF134674
		Glycosmis	Glycosmis trichanthera	A,B,E	EF118879	AF320855	EF138894	EF164866	EF176550	EF126616	EF126683	EF134686
			Glycosmis mauritiana	A,B,E	EF118878	AF320862	EF138892	EF164864	EF176548	EF126614	EF126681	EF134684
			Glycosmis pentaphylla	A,B,E	EF118877	EF126548	EF138893	EF164865	EF176549	EF126615	EF126682	EF134685
		Murraya	Murraya koenigii	C,D,E	EF118832	AF320867	EF138843	EF164815	EF176499	AF320262	EF126637	EF134635
			Murraya paniculata	F, G	EF118891	AF320868	EF138906	EF164878	EF176562	AY295254	AY295280	EF134698
	Merrilliinae	Merrillia	Merrillia caloxylon	F, G	EF118882	AF320871	EF138897	EF164869	EF176553	AY295270	AY295296	EF134689
Citreae	Triphasiinae	Luvunga	Luvunga	J,H,G	EF118880	EF126549	EF138895	EF164867	EF176551	EF126617	EF126684	EF134687
		Merope	Merope angulata	L,I,H,G	EF118881	EF126550	EF138896	EF164868	EF176552	EF126618	EF126685	EF134688
		Monanthocitrus	Monanthocitrus cornuta	L,I,H,G	EF118890	EF126557	EF138905	EF164877	EF176561	EF126624	EF126692	EF134697
		Oxanthera	Oxanthera neocaledonica	OO, TT, T, R, Q, O, M, H, G	EF118893	EF126559	EF138908	EF164880	EF176564	EF126625	EF126693	EF134700
			Oxanthera sp	OO, TT, T, R, Q, O, M, H, G	EF118894	EF126560	EF138909	EF164881	EF176565	EF126626	EF126694	EF134701
		Pamburus	Pamburus missionis	J, H, G	EF118895	AF320883	EF138910	EF164882	EF176566	AY295274	AY295300	EF134702
		Paramignya	Paramignya lobata	J, H, G	EF118896	EF126561	EF138911	EF164883	EF176567	EF126627	EF126695	EF134703
			Paramignya scandens	J, H, G	EF118897	EF126562	EF138912	EF164884	EF176568	AY295257	EF126696	EF134704
		Triphasia	Triphasia trifolia	L,I,H,G	EF118902	AF320884	EF138921	EF164893	EF176577	AY295271	AY295297	EF134713
		Wenzelia	Wenzelia dolichophylla	L,I,H,G	EF118903	EF126566	EF138922	EF164894	EF176578	AY295260	AY295286	EF134714
	Balsamocitrinae	Aegle	Aegle marmelos	K, H, G	AF066839	AF320882	EF138836	EF164808	EF176492	AY295268	AY295294	EF134628
		Aeglopsis	Aeglopsis chevalieri	K, H, G	EF118827	EF126500	EF138837	EF164809	EF176493	EF126567	EF126634	EF134629
		Afraegle	Afraegle paniculata	K, H, G	EF118828	EF126501	EF138838	EF164810	EF176494	AY295269	AY295295	EF134630
		Balsamocitrus	Balsamocitrus dawei	K, H, G	EF118831	EF126504	EF138842	EF164814	EF176498	EF126571	AY295278	EF134634
		Feronia	Feronia limonia	P, O, M, H, G	EF118870	EF126541	EF138885	EF164857	EF176541	EF126609	EF126676	EF134677
		Feroniella	Feroniella oblata	CC, BB, Y, SS, S, Q, O, M, H, G	EF118871	EF126542	EF138886	EF164858	EF176542	AY295263	AY295289	EF134678
		Swinglea	Swinglea glutinosa	N, M, H, G	EF118899	AF320887	EF138918	EF164890	EF176574	AY295259	EF126699	EF134710

Citrinae	Fruit proche des agrumes	Atalantia	Atalantia ceylanica	P, O, M, H, G	AF066840	EF126502	EF138839	EF164811	EF176495	EF126568	AY295288	EF134631
			Atalantia citroides	P, O, M, H, G	EF118829	EF126503	EF138840	EF164812	EF176496	EF126569	EF126635	EF134632
			Atalantia monophylla	P, O, M, H, G	EF118830	AF320874	EF138841	EF164813	EF176497	EF126570	EF126636	EF134633
		Citropsis	Citropsis daweanana	N, M, H, G	EF118837	EF126509	EF138848	EF164820	EF176504	EF126576	EF126642	EF134640
			Citropsis schweinfurthii	N, M, H, G	EF118838	EF126510	EF138849	EF164821	EF176505	EF126577	EF126643	EF134641
	Fruit primitif	Hesperethusa	Hesperethusa crenulata	N, M, H, G	EF118892	EF126558	EF138907	EF164879	EF176563	AY295272	AY295298	EF134699
		Pleiospermium	Pleiospermium latialatum	N, M, H, G	EF118898	EF126563	EF138913	EF164885	EF176569	EF126628	EF126697	EF134705
		Severinia	Severinia buxifolia	P, O, M, H, G	AF066835	AF320886	EF138916	EF164888	EF176572	EF126629	AY295290	EF134708
		Burkillanthus	Burkillanthus malaccensis	N, M, H, G	EF118833	EF126505	EF138844	EF164816	EF176500	EF126572	EF126638	EF134636
	Agrume vrai	Citrus	C. maxima	GG4, EE, Z, Y, SS, S, Q, O, M, H, G	EF118839	EF126529	EF138869	EF164841	EF176525	EF126597	EF126663	EF134661
			C. medica	QQ, R, Q, O, M, H, G	EF118860	EF126531	EF138871	EF164843	EF176527	EF126599	EF126665	EF134663
			C. ichangensis	JJ, II, CC, BB, y, SS, S, Q, O, M, H, G	EF118852	EF126522	EF138862	EF164834	EF176518	EF126590	EF126656	EF134654
				HH, BB, Y, SS, S, Q, O, M, H, G	EF118849	EF126519	EF138859	EF164831	EF176515	EF126587	EF126653	EF134651
			C. reticulata	JJ, II, CC, BB, y, SS, S, Q, O, M, H, G	EF118864	EF126537	EF138877	EF164849	EF176533	EF126604	EF126670	EF134669
				MM, AA, X, SS, S, Q, O, M, H, G	EF118858	EF126528	EF138868	EF164840	EF176524	EF126596	EF126662	EF134660
			C. aurantifolia	L2, FF1, DD, Z, Y, SS, S, Q, O, M, H, G	EF118841	AF320875	EF138851	EF164823	EF176507	EF126579	EF126645	EF134643
				GG1, EE, Z, Y, SS, S, Q, O, M, H, G	EF118853	EF126523	EF138863	EF164835	EF176519	EF126591	EF126657	EF134655
			C. aurantium	GG1, EE, Z, Y, SS, S, Q, O, M, H, G	EF118842	EF126512	EF138852	EF164824	EF176508	EF126580	EF126646	EF134645
				GG2, EE, Z, Y, SS, S, Q, O, M, H, G	EF118843	EF126513	EF138853	EF164825	EF176509	EF126581	EF126647	EF134644
				EE, Z, Y, SS, S, Q, O, M, H, G	EF118863	EF126536	EF138876	EF164848	EF176532	EF126602	EF126669	EF134667
				GG1, EE, Z, Y, SS, S, Q, O, M, H, G	EF118861	EF126532	EF138872	EF164844	EF176528	EF126600	EF126666	EF134664
			C. limon	JJ, II, CC, BB, y, SS, S, Q, O, M, H, G	EF118856	EF126526	EF138866	EF164838	EF176522	EF126594	EF126660	EF134658

	GG1, EE, Z, Y, SS, S, Q, O, M, H, G	EF118855	EF126525	EF138865	EF164837	EF176521	EF126593	EF126659	EF134657
	JJ, II, CC, BB, Y, SS, S, Q, O, M, H, G	EF118851	EF126521	EF138861	EF164833	EF176517	EF126589	EF126655	EF134653
C. paradisis	EE, Z, Y, SS, S, Q, O, M, H, G	EF118862	EF126533	EF138873	EF164845	EF176529	EF126601	EF126667	EF134665
	GG5, EE, Z, Y, SS, S, Q, O, M, H, G	AJ238408	EF126534	EF138874	EF164846	EF176530	AY295251	AY295277	EF134666
C. sinensis	GG6, EE, Z, Y, SS, S, Q, O, M, H, G	EF118866	EF126539	EF138879	EF164851	EF176535	EF126606	EF126672	EF134671
C.amboiensis	L1, FF1, DD, Z, Y, SS, S, Q, O, M, H, G	EF118840	EF126511	EF138850	EF164822	EF176506	EF126578	EF126644	EF134642
C.halimii	LL, X, SS, S, Q, O, M, H, G	EF118847	EF126517	EF138857	EF164829	EF176513	EF126585	EF126651	EF134649
C.hystrix	L1, FF1, DD, Z, Y, SS, S, Q, O, M, H, G	EF118848	EF126518	EF138858	EF164830	EF176514	EF126586	EF126652	EF134650
	L3, FF1, DD, Z, Y, SS, S, Q, O, M, H, G	EF118845	EF126515	EF138855	EF164827	EF176511	EF126583	EF126649	EF134647
	FF2, FF1, DD, Z, Y, SS, S, Q, O, M, H, G	EF118857	EF126527	EF138867	EF164839	EF176523	EF126595	EF126661	EF134659
C.indica Tanaka	QQ, R, Q, O, M, H, G	EF118850	EF126520	EF138860	EF164832	EF176516	EF126588	EF126654	EF134652
C.latipes (Swing.) Tan.	GG3, EE, Z, Y, SS, S, Q, O, M, H, G	EF118854	EF126524	EF138864	EF164836	EF176520	EF126592	EF126658	EF134656
C.tachibana (Mak.) Tan.	KK, II, CC, BB, Y, SS, S, Q, O, M, H, G	EF118867	EF126540	EF138880	EF164852	EF176536	EF126607	EF126673	EF134672
C.tangelo	EE, Z, Y, SS, S, Q, O, M, H, G	EF118834	EF126535	EF138875	EF164847	EF176531	EF126603	EF126668	EF134668
Citrofortunella floridana	L1, FF1, DD, Z, Y, SS, S, Q, O, M, H, G	EF118836	EF126508	EF138847	EF164819	EF176503	EF126575	EF126641	EF134639
C.maxima x C. medica	GG1, EE, Z, Y, SS, S, Q, O, M, H, G	EF118859	EF126530	EF138870	EF164842	EF176526	EF126598	EF126664	EF134662
C.reticulata 'fina' x C. glauca (eremocitrus glauca)	JJ, II, CC, BB, Y, SS, S, Q, O, M, H, G	EF118865	EF126538	EF138878	EF164850	EF176534	EF126605	EF126671	EF134670

				Citrus gracilis	W, U, TT, T, R, Q, O, M, H, G	EF118846.1	EF126516.1	EF138856.1	EF164828.1	EF176512.1	EF126584.1	EF126650.1	EF134648
				Citrus x oliveri	V, U, TT, T, R, Q, O, M, H, G	EF118876.1	EF126547.1	EF138891.1	EF164863.1	EF176547.1	EF126613.1	EF126678.1	EF134683
		Clymenia	Clymenia polyandra	PP, T, R, Q, O, M, H, G		EF118869	AF320878	EF138883	EF164855	EF176539	AY295255	AY295281	EF134675
		Eremocitrus	Eremocitrus glauca	Eremocitrus, W, U, TT, T, R, Q, O, M, H, G		AF066847	AF320879	EF138884	EF164856	EF176540	AY295267	AY295293	EF134676
		Fortunella	Fortunella japonica	NN, AA, X, SS, S, Q, O, M, H, G		EF118873	EF126544	EF138888	EF164860	EF176544	EF126611	EF126679	EF134680
			Fortunella margarita	NN, AA, X, SS, S, Q, O, M, H, G		EF118874	EF126545	EF138889	EF164861	EF176545	EF126612	EF126680	EF134681
			Fortunella polyandra	MM, AA, X, SS, S, Q, O, M, H, G		EF118875	EF126546	EF138890	EF164862	EF176546	AY295265	AY295291	EF134682
		Microcitrus	Microcitrus australis	W, U, TT, T, R, Q, O, M, H, G		EF118884	EF126552	EF138899	EF164871	EF176555	EF126620	EF126687	EF134691
			Microcitrus australasica	V, U, TT, T, R, Q, O, M, H, G		EF118883	EF126551	EF138898	EF164870	EF176554	EF126619	EF126686	EF134690
			Microcitrus inodora	W, U, TT, T, R, Q, O, M, H, G		EF118886	EF126554	EF138901	EF164873	EF176557	EF126621	EF126688	EF134693
			Microcitrus garrawayi	V, U, TT, T, R, Q, O, M, H, G		EF118885	EF126553	EF138900	EF164872	EF176556	AY295261	AY295287	EF134692
			Microcitrus papuana	W, U, TT, T, R, Q, O, M, H, G		EF118887	EF126555	EF138902	EF164874	EF176558	EF126622	EF126689	EF134694
			Microcitrus warburgiana	W, U, TT, T, R, Q, O, M, H, G		EF118888	EF126556	EF138903	EF164875	EF176559	EF126623	EF126690	EF134695
		Poncirus	Poncirus trifoliata	RR, S, Q, O, M, H, G		AJ238409	AF320876	EF138914	EF164886	EF176570	AY295256	AY295282	EF134706

Annexe 2 : Liste du matériel végétal utilisé dans l'analyse KASPar

Tribu	Sous-tribu		Genre	Espèce	Code	Collection
Clauseneae	Clauseniae		Clausena	Clausena anisata (Willd.)	IVIA 347	IVIA
				Clausena excavata Burm.	IVIA 311	IVIA
				Clausena lansium (Lour.)	IVIA 343	IVIA
			Glycosmis	Glycosmis pentaphylla (Retz.)	IVIA 148	IVIA
			Murraya	Murraya koenigii (L.)	IVIA 377	IVIA
				Murraya paniculata (L.)	IVIA 283	IVIA
	Merrilliinae		Merrillia	Merrillia cloxylon (Ridl.) Swing.	RCRC 3786	USDA-ARS
Citreae	Triphasilineae		Oxanthera	Oxanthera neo-caledonica (Guillaumin) Tanaka	RCRC 3793	USDA-ARS
			Pamburus	Pamburus missionis (Wall. Ex Wight) SwinG.	RCRC 3511	USDA-ARS
			Paramignya	Paramignya scandens (Griff.) Craib	PI 109758	USDA-ARS
			Triphasia	Triphasia trifolia (Burm. F.)	IVIA 182	IVIA
			Wenzelia	Wenzelia dolichophylla (Lauterb and K.Schum.) Tanaka	RCRC 3966	USDA-ARS
	Balsamocitrinae		Aegle	Aegle marmelos (L.)	IVIA 345	IVIA
			Aeglopsis	Aeglopsis chevalieri Swing.	IVIA 308	IVIA
			Afraegle	Afraegle gabonensis	RCRC 4191	USDA-ARS
				Afraegle paniculata (Schum.)	IVIA 273	IVIA
			Balsamocitrus	Balsamocitrus dawei Stapf	IVIA 372	IVIA
			Feronia	Limonia acidissima (Feronia limonia)	RCRC 3299	USDA-ARS
			Feroniella	Feroniella oblata Swing	IVIA 585	IVIA
			Swinglea	Swinglea glutinosa (Blanco)	IVIA 292	IVIA
	Citrinae	Fruit proche des agrumes	Atalantia	Atalantia ceylanica (Arn.)	IVIA 172	IVIA
				Atalantia citroides Pierre ex Guill.	IVIA 180	IVIA
				Atalantia monophylla (L.) DC	RCRC 4189	USDA-ARS
				Atalantia roxburghiana	RCRC 4190	USDA-ARS
			Citropsis	Citropsis daweani Swingle and M.kell	RCRC 3294	USDA-ARS
				Citropsis gabunensis	RCRC 3286	USDA-ARS
				Citropsis gillettiana Swing. and M.Kell	IVIA 517	IVIA

			Citropsis schweinfurthii (Engl.) Swing. and M. Kell.	IVIA 310	IVIA
		Hesperethusa	Hesperethusa crenulata (Roxb.)	IVIA 580	IVIA
		Pleiospermium	Pleiospermium alatum Swingle	RCRC 3111	USDA-ARS
		Severinia	Severinia buxifolia (Poir.)	IVIA 147	IVIA
			Severinia disticha (Blanco) Swing.	IVIA 418	IVIA
			C. maxima (Burm.) Osbeck	IVIA 589	IVIA
				SRA 608	INRA-CIRAD
				SRA 707	INRA-CIRAD
				SRA 757	INRA-CIRAD
				SRA 321	INRA-CIRAD
				IVIA 590	IVIA
				SRA 322	INRA-CIRAD
				SRA 727	INRA-CIRAD
			C. medica (L.)	SRA613	INRA-CIRAD
				SRA640	INRA-CIRAD
				SRA130	INRA-CIRAD
				SRA722	INRA-CIRAD
				SRA760	INRA-CIRAD
				SRA701	INRA-CIRAD
			C. ichangensis Swing.	IVIA 358	IVIA
				SRA 0110062	INRA-CIRAD
			C. reticulata Blanco	IVIA478	IVIA
				IVIA 135	IVIA
				IVIA 359	IVIA
				IVIA 238	IVIA
				IVIA22	IVIA
				SRA 839	INRA-CIRAD
				SRA 769	INRA-CIRAD
				SRA 586	INRA-CIRAD

163

				C. sinensis (L.)	SRA 246	INRA-CIRAD
					SRA 203	INRA-CIRAD
				C.hystrix DC.	SRA630	INRA-CIRAD
				C.halimi B.C Stone	IVIA 278	IVIA
				C.micrantha Wester.	SRA1044	INRA-CIRAD
			Clymenia	Clymenia polyandra (Tan.) Swing.	RCRC 3284	USDA-ARS
			Eremocitrus	Eremocitrus glauca (Lindl.) Swing.	IVIA 346	IVIA
			Fortunella	Fortunella hindsii (Champ.) Swing.	IVIA 280	IVIA
				Fortunella crassifolia Swing.	IVIA 281	IVIA
				Fortunella japonica (Thumb.) Swing.	IVIA 381	IVIA
				Fortunella margarita (Lour.) Swing.	IVIA 38	IVIA
				Fortunella obovata Tan.	IVIA 312	IVIA
				Fortunella polyandra (Ridl.) Tan.	IVIA 375	IVIA
			Microcitrus	Microcitrus australis (Planch.) Swing.	IVIA 150	IVIA
				Microcitrus australasica (F. Muell.) Swing.	IVIA 313	IVIA
				Microcitrus inodora (F.M. Bail.) Swing.	RSD 2002021	USDA-ARS
				Microcitrus garrowayae (F. M.) Bailey	IVIA 314	IVIA
				Microcitrus papuana Mabb.	RCRC 4111	
				Microcitrus virgata Mabb.	IVIA 315	IVIA
				Microcitrus warburgiana (F.M. Bail.) Tan.	RCRC 1466	USDA-ARS
				M. australisica x C. mitis	RCRC 1485	USDA-ARS
			Poncirus	Poncirus trifoliata (L.) Raf.	IVIA 374	IVIA
					IVIA 217	IVIA

Annexe 3 : Liste des SNPs sélectionnés et leurs séquences avec 50 pb de chaque côté du SNP

Marqueur	Position	Clade	Séquence
CP_58345	58345	QQ	ATATATATACCCTGCCTTTCTTTACTGGATGAATTACGCATATTTTCTTT[G/A]CACATATAGGATTTACGTATACAACACATATTACTGTCAAGAGTCAATTT
CP_4296	4296	QQ	GCGATTGAACAAGAATAAAATATTCTCAAAATTATCCATTGATACGACAT[G/A]CTGTTTTTTCCATTCAATTCCTTTTCAGGATCAGTCGCAGTCTTACAACTC
CP_32299	32299	R	TTGGCAATACGCTTAAAAAAGATTATTCATTCCCTCTTCGGGAGGGGTTG[G/T]CGCCTTGCTTGATCTTTATTTTATTAAGAATATCATCCCCCTCCCTTGG
CP_58401	58401	S	ATAGGATTTACGTATACAACACATATTACTGTCAAGAGTCAATTTCTTAT[G/T]GATTGGGTTGCGCCATACATATGAAAGAGTATACAATAATGATGATTTTG
CP_56697	56697	D	AGCTATAATGTCCTGAAGTTCTTTGTAACGTTGTAAGTTTGCTTAACCT[C/T]TTGCGCCGTTTCATAATGTTCTCGCCAACGATCCGAGGTTGCAGCATAG
CP_3807	3807	E	AAGTATTGTAGTAAAAAAGTGGATACCTTGAAATGGGTAGACTCATCAC[C/T]GGATTCTCTATCCTCTCATTTTGAGTTAATTTAATTGGTTTATGTTTGT
CP_58454	58454	G	TTAGAAATTTGCGATTGAAAAAGGTAAGGGATTAGAACTTGAAAGGCGC[C/T]GATTGGGTTGCGCCATACATATGAAAGAGTATACAATAATGATGATTTTG
CP_32526	32526	H	AATCTTAAAGTATTCTATCAAAAGTAACGATCTTAAATTCAGTTGTGTAT[A/G]CGCTCCCGGGAACGATATGGTACTCGATTCCATTCCATACACAGATGGGG
CP_3964	3964	Eremocitrus	CTTTTGATTTTGAAAAAACAACACTATTTTATCAATATACTGCTTCTT[C/A]TACACATTCAGTTACAACCCATAATAGGGACCCGCTAATACTTAGGACTC
CP_9149	9149	LL	AGTATTCGAAAGAGATAAAATAGAAAAAATAAAGAAAAGAATATCTCTT[T/A]AATTCGTTCAACGGGGCCTTTTTTATTTCCACTTCCACGGCCTGGCCTG
CP_58062	58062	X	CAATAGTTTCCTTATGCTTATGCATTTATGCAATTTCAATGAGGGAATTT[T/G]CAAGTTCAACCAACCGACTTTCAAATATCAAGTGGGTGAATAAAAATCT
CP_10342	10343	AA	CCCGGCTTCTCTCTGGGAATATAGCAAACCTCCATAAAGGAGCCGAATGA[C/A]CCCAAAGTTTCATGTTTCGGTTTTGAATTAGCGGCGTTAAAAAGTATGAAT
CP_51634	51634	FF	CATACTGAAACTTACAAAGTCTTCTTTTTGATGATTCAAGAAATGAAATT[C/A]CCCTCCCAAGACTTTTAAATCCCTGTTATTTTTTAATTGACATAGACCCA
CP_3674	3674	V	TAGGAAGTCGTGTTGTTGAGATCTATCTAGTTCTAAATATACTTGAAATT[C/G]CTCCATTTGAAATTCGATTAAAAACAAAGGTAAGGGATTTAGTGAGCGAT
CP_4243	4243	W	TTAATAATTATATTTATTTTGTTCGCGCCAAGAATCAAACCTTGGTTTT[A/C]TAGCGATTGAACAAGAATAAAATATTCTCAAATATCCATTGATACGAC
CP_57691	57691	RR	CTAACTCGATTATTCCTAATAATTGCTGTACCTCGCAAGTCACATTAAT[T/C]GGTTGATCGACAGTATCTCGACCCTTAACCTACCAGAGCATTGTAAATATT
CP_4159	4159	CC	AATTAAGAACCGAAGCTCGTTACTTTTTGTTTCCCTATAATTGGAACCCCT[A/C]GGGCTCTATCCATTTATTAACCTCGACCCAACCTTAATAATTATATTTAT
CP_57064	57064	F	TCGGATCCCGCTTGACGAAACGGAAGATATTGTCGATAAATAGAAGTAC[G/A]TCTTGTTCAATAACATCTCGGAAATATTCCGCCATAGTTAGGGCCGTC
CP_32463	32458	J	AATTCAACTAGTAATTTAAGTTAGCCAACTAGAAACCAGAAAAGAAGA[T/C]ACTTTGAATCTTAATCTTAAAGTATTCTATCAAAGTAACGATCTTAAAT
CP_9158	9158	K	AAGAGATAAAATAGAAAAAATAAAGAAAAGAATATCTCTTAAATTCGT[T/C]CAACGGGGCCTTTTTTATTTCCACTTCCACGGCCTGGCCTGGTCAGTACC
CP_3829	3829	M	ATACCTTGAAATGGGTAGACTCATCACCGGATTCTCTATCCTCTCATTT[T/C]GAGTTAATTTAATTGGTTTATGTTTGTATAGTTATAGTATAACTAAGTG
CP_49405	49405	M	TCCAGACCCCTTTTTGTATTGTATACGCATATAGAAATATAAATCAAAG[A/G]GGGTTCTTTTCTTGATTATTCTGCCGAAATCGAACTTCCTCCCACTTT
CP_3747	3747	N	AACAAAGGTAAGGGATTTAGTGAGCGATCAAACGATACATAGTGCGATAC[G/A]GTGAAACAAAGTATTGTAGTAAAAAGTGGATACCTTGAAATGGGTAG
CP_51068	51068	OO	GTATATAGACTATGTATACGCAGCGAAAACTAACTCAAAAATCACAACC[A/C]AATACGTATTTCTTTTATGAAAAAGAAAAGAAATTGTTATTGTTATGAAT
CP_3982	3982	P	AAACAACCTATATTTATCAATATACTGCTTCTTCTACACATTCAGTTACAA[C/T]CCATAATAGGGACCCGCTAATACTTAGGACTCATTAATAAATCGATAAT

CP_49301	49301	Q	ATTGAATTTTCATAAACCTAAATTTAACTGAGCTACATGATAAATGAAGC[A/G]AAATCCGCTGAAGTATTGTACTACAAAGAATACTAAGATGATTTGTATCA
CP_58553	58553	Q	GATGTATTTGGTGAATCAAATACCATGGTCTAAAAAACAAGGAAACGCTCTGAGTAG[G/T]TGATAATATTAATTGAAAAGCTTGAAAAGATTCTGTGAAGG GGTTTCATTAACCTCCTAAT
CP_51032	51032	II	CGAATATCGAACTCGATAAAGGATGAAGGATAAGGGTATATAGACTATG[T/G]ATACGCAGCGAAAACTAACTCAAAAATCACAACCAAATACGTATTTCTT
CP_57686	57686	HH	CAGCTCTAACTCGATTATTCCTTAATAATTGCTGTACCTCGCAAGTCACA[T/C]TAATTGGTTGATCGACAGTATCTCGACCCTTAACCTACCAGAGCATTGTAA
CP_3707	3707	DD	TAAATATACTTGAAATTCCTCCATTTGAAATTCGATTAAAAACAAAGGTA[A/C]GGGATTTAGTGAGCGATCAAACGATACATAGTGCGATACGGTGAAAACAA
CP_10192	10193	U	AATAGAAATAGAGAACGAAGTAAGTAGAAAGATTGGTATTGTTATAATCC[C/G]ACTCTTCTAGAGGGATCATCTAGAAAGCGGATTGTTTTGAATTCATTCAG
CP_32282	32277	N	TTAGAGTTTCTTCTTCTTGGCAATACGCTTAAAAAAGATTATTCATTCC[C/G]TCTTCGGGAGGGGTTGGCGCCTTGCTTGATCTTTATTTTATTAAGAATAT
CP_5927	5927	EE	TTACTTACGAAGTTTTTCCAACCTATTGATTGGTATTAACCCTAGATCC[G/T]TGCCCCTGAGAAAGGAATAAATACTTTTACTCGAGCTCCATCCTGTACT
CP_32684	32679	Z	ACCAATAATTGTAGCAATTCGCACATGTCTCTTCTCATCAATCGGTACT[A/G]GTTGATGAGAAATCGAAATGAAAAAGAAAAAGAGCAAAGGAGAGCAGTA
CP_57022	57022	A	TAACCCACAGCGGAAGGCATTCTACCCAATAAGGCGGATACCTCGGATCC[C/T]GCTTGGACGAAACGGAAGATATTGTCGATAAATAGAAGTACGTCTTGTTG
CP_58121	58121	J	ACCAACCGACTTTCAAATATCAAGTGGGTGAATAAAAAATCTTGAGAAAG[T/C]CGTTTATTATCTATCATTATAGACAATACCCGCCATATTATCTATGGA
CP_51540	51540	JJ	ATCACAAGCCGTATGGTCTATACGATATATGTAGAAATGAACACCTTTGA[G/T]CAAGGAATCCCCGTTTTAATGATTCCCAATCCATATTATTGCTCATACTG
CP_57256	57256	L	CCAGACTCTTTCATTTCCATGTAAAGATCATTTCTTCACGAGTACGTT[C/A/C]CCGACTCCGCCAAATACGGATACACCCCATGGGCTTTGGCAATGTTGTT
CP_57742	57742	PP	GGTTGATCGACAGTATCTCGACCCTTAACCTACCAGAGCATTGTAAATATT[A/C]GGCATCTTGCCGGGGGGAAAGGCTACATCCAGTACCGGACCAATGATTTG
CP_3996	3996	Q	ATCAATATACTGCTTCTTCTACACATTCAGTTACAACCCATAATAGGGAC[C/T]CGCTAATACTTAGGACTCATTAAATAAATCGATAATCCCCTCATGGGAAA

[illegible]

Annexe 5 : Liste du matériel végétal utilisé

Commun name	Horticultural group	Parents	Latin name (Swingle and Reece, 1967)	SRA Ref
Chandler pummelo	pummelo		<i>Citrus maxima</i> (Burm.) Merr	SRA 608
Deep red pummelo	pummelo		<i>Citrus maxima</i> (Burm.) Merr	SRA 757
KaoPan pummelo	pummelo		<i>Citrus maxima</i> (Burm.) Merr	SRA 321
Pink pummelo	pummelo		<i>Citrus maxima</i> (Burm.) Merr	SRA 322
Tahiti pummelo	pummelo		<i>Citrus maxima</i> (Burm.) Merr	SRA 727
Timor pummelo	pummelo		<i>Citrus maxima</i> (Burm.) Merr	SRA 707
Chios mandarin	mandarin		<i>Citrus reticulata</i> Blanco	ICVN 0100598
Cleopatra mandarin	mandarin		<i>Citrus reticulata</i> Blanco	ICVN 0110066
Dancy mandarin	mandarin		<i>Citrus reticulata</i> Blanco	SRA 594
Fuzhu mandarin	mandarin		<i>Citrus reticulata</i> Blanco	ICVN 0100775
King mandarin	mandarin		<i>Citrus reticulata</i> Blanco	SRA 583
Ladu mandarin	mandarin		<i>Citrus reticulata</i> Blanco	SRA 595
Ponkan mandarin	mandarin		<i>Citrus reticulata</i> Blanco	SRA 584
San Hu Hong Chu mandarin	mandarin		<i>Citrus reticulata</i> Blanco	SRA 769
Satsuma Owari mandarin	mandarin		<i>Citrus reticulata</i> Blanco	SRA 221
Sunki mandarin	mandarin		<i>Citrus reticulata</i> Blanco	ICVN 0110078
Szibat mandarin	mandarin		<i>Citrus reticulata</i> Blanco	SRA 596
Carvalhal mandarin hybrid	tangor	hyb naturel	<i>Citrus reticulata</i> Blanco	SRA 271
Clemenules clementine	tangor	hyb naturel	<i>Citrus reticulata</i> Blanco	ICVN 0100389
Dweet tangor	tangor	Dancy x Sweet orange	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	ICVN0100464
Ellendale tangor	tangor	hyb naturel	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 656
Kiyomi tangor	tangor	Miyagawa-wase (Satsuma mandarin) X Trovita orange	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 704
Murcott tangor	tangor	hyb naturel	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	ICVN 0100601
Ortanique tangor	tangor	hyb naturel	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 110
Temple tangor	tangor	hyb naturel	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 280
Allspice tangelo	tangelo	Imperial pomelo x Willow leaf	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 327
Mapo tangelo	tangelo	Duncan x Willow leaf	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 450
Minneola tangelo	tangelo	Duncan x Dancy	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 451
Orlando tangelo	tangelo	Duncan x Dancy	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 46
Pearl tangelo	tangelo	Willow leaf x Imperial pomelo	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 296
Sampson tangelo	tangelo	pomelo x dancy	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	ICVN0100456

Annexe 5 : Liste du matériel végétal utilisé (suite)

Commun name	Horticultural group	Parents	Latin name (Swingle and Reece, 1967)	SRA Ref
San Jacinto tangelo	tangelo	hyb naturel	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 297
Seminole tangelo	tangelo	Duncan x Dancy	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 298
Sunrise tangelo	tangelo	Duncan x Dancy	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 457
Sunshine tangelo	tangelo	Duncan x Dancy	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	ICVN0100458
UGLI® tangelo	tangelo	hyb naturel	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 950
Webber SG tangelo	tangelo	Duncan x Dancy	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	ICVN0100475
Fairchild tangelo	tangor x tangelo	Clem x orlando	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 30
Nova tangelo	tangor x tangelo	clem x orlando	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 158
Osceola tangelo	tangor x tangelo	clem x orlando	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 48
Page tangelo	tangor x tangelo	clem x minneola	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 159
Robinson tangelo	tangor x tangelo	clem x orlando	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 47
Fallglo mandarin	(tangor x tangelo) x tangor	(ClemxOrlando)x Temple	<i>Citrus reticulata</i> Blanco	IVIA 466
Kara mandarin	Mand hybrid	King x Satsuma	<i>Citrus reticulata</i> Blanco	SRA241
Wilkling mandarin	Mand hybrid	Willow leaf x King	<i>Citrus reticulata</i> Blanco	SRA112
Jackson orangelo	orangelo	hyb naturel	<i>Citrus sinensis</i> (L.) Osbeck x <i>Citrus paradisi</i> Macfad.	SRA 472
Triumph orangelo	orangelo	hyb naturel	<i>Citrus sinensis</i> (L.) Osbeck x <i>Citrus paradisi</i> Macfad.	ICVN0100474
Wekiwa SG tangelo	pomelo x tangelo	sampson tangelo x pomelo	<i>Citrus reticulata</i> Blanco x <i>Citrus paradisi</i> Macfad.	SRA 462
Fortune mandarin	Tangor x Mand	Clem x Dancy	<i>Citrus reticulata</i> Blanco	SRA 31
Fremont mandarin	Tangor x Mand	Clem x Ponkan	<i>Citrus reticulata</i> Blanco	SRA 147
Ambersweet tangor	tangor x tangelo x orange	hybride complexe	<i>Citrus reticulata</i> Blanco x <i>Citrus sinensis</i> (L.) Osbeck	SRA 875
Marsh grapefruit	grapefruit		<i>Citrus paradisi</i> Macfad.	SRA 188
Star Ruby grapefruit	grapefruit		<i>Citrus paradisi</i> Macfad.	ICVN 0100293
Bigaradier Maroc sour orange	sour orange		<i>Citrus aurantium</i> L.	ICVN0110033
Washington Navel Sweet orange	sweet orange		<i>Citrus sinensis</i> (L.) Osbeck	SRA 203

Annexe 6 : Les régions d'introgressions chez les variétés représentatives des taxons de référence *Citrus reticulata* et *Citrus maxima*

Chromosome	Variété	Région introgressé	
		Début	Fin
1	Satsuma	27 597 556.6	28 931 563
	King	28 233 222.1	
	KaoPan	25 111 191.5	
	Tahiti	25 111 191.5	
2	Satsuma	11 473 499.7	22 843 147.7
	Hong chu	28 318 800.3	31 781 542.6
	Szibat	29 092 870.2	36 357 909
	Dancy	29 092 871.2	36 357 909
	Fuzhu	29 092 872.2	36 357 909
	Ladu	31 781 542.6	33 629 465.3
	Ponkan	29 092 872.2	36 357 909
	King	25 455 930.9	36 357 909
	Chandler	9 615 659.6	
		24 490 037.8	25 926 762.2
	Kaopan	24 490 037.8	25 926 762.2
	Tahiti	34 917 476.2	35 080 942.5
3	Satsuma	17 309	46 228 541.3
		51 012 039.6	51 024 513
	Chios	17 309	5 957 353.0
	Dancy	17 309	5 809 191.1
	Cleopatre	2 885 087.8	5 580 895.9
	Fuzhu	17 309	6 242 340.5
	Ladu	2 150 000.0	6 242 340.5
	Ponkan	17 309	5 957 353.0
	Sunki	17 309	2 051 652.2
		3 100 000.0	6 242 340.5
		28 677 279.3	29 546 735.7
	King	17 309	10 808 380.6
		25 139 086.8	27 426 049.2
		31 272 475.5	36 830 463.7
		42 897 127.8	47 500 094.4
4	Chios	6 477 832.2	9 873 182.7
	Szibat	1 551 466.4	2 598 230.7
	King	22 982 500.8	25 629 994
5	Satsuma	12 876 429.3	19 113 595.85
		33 769 487.6	43 289 355
	King	65 072	5 161 688.0

Annexe 6 : Les régions d'introgressions chez les variétés représentatives des taxons de référence *Citrus reticulata* et *Citrus maxima* (Suite)

Chromosome	Variété	Région introgressée	
		Début	Fin
6	Chios	18 688 093.0	25 582 311
	Satsuma	46 824	3 147 962.7
		5 293 341.5	13 262 850.1
		22 447 865.2	22 597 520.8
	Fuzhu	18 990 071.7	22 387 017.5
	Szibat	21 697 195.1	22 227 617.2
	Dancy	18 990 071.7	22 387 017.5
	Ponkan	18 990 071.7	25 582 311
	King	46 824	3 147 962.7
		5 293 342.5	13 625 933.4
		14 262 429.1	25 097 424.1
7	Satsuma	31 028	1 075 177.7
		1 295 775.3	6 526 585.9
		6 637 230.9	21 091 285
	Cléopâtre	2 938 108.6	3 128 094.0
	King	31 028	8 317 564.6
8	Chios	17 942	5 499 396
		18 012 803	22 198 462
	Satsuma	17 942	23 698 043
	Ladu	17 942	2 817 597.7
		18 799 669.9	22 408 491.9
	King	17 942	25 110 297
	Sunki	7 000 000.00	7 400 000.00
	San Hu Hong Chu	17 942	3 825 531
		16 179 594.7	17 478 687.3
	Md Szibat	17 942	1 613 564.6
	Dancy	17 942	2 283 351.7
		18 799 669.9	22 400 000.0
	Fuzhu	17 942	2 600 000.0
		18 799 669.9	22 408 491.9
	Ponkan	17 942	2 817 597.7
		16 179 594.65	16 880 372.4
9	Satsuma	22 625	450 052.0
		10 863 652.2	14 922 737.0
	Hong Chu	7 749 253.3	9 916 568.8
		14 922 737.0	15 396 677.0
		15 639 915.2	28 120 177.8
	Szibat	31 250 334.5	31 263 015
	Ponkan	4 690 202.3	10 863 652.2
		15 211 961.2	28 120 177.8
	King	22 625	2 282 546.2
		4 943 901.5	28 120 177.8
	Deep red	3 034 987.5	3 738 741.6
	Timor	4 570 516.0	5 690 202.3

Annexe 7 : Analyse de l'hétérozygotie observée et estimation des données manquantes chez les variétés étudiées au niveau de chaque chromosome

	Missing data										Ho									
	Chr1	chr2	Chr3	Chr4	Chr5	Chr6	Chr7	Chr8	Chr9	Total	Chr1	chr2	Chr3	Chr4	Chr5	Chr6	Chr7	Chr8	Chr9	Total
Chandler pummelo	9.16%	7.54%	8.13%	8.72%	11.06%	8.57%	7.86%	11.06%	9.07%	9.01%	0.00%	2.10%	0.00%	0.00%	0.00%	0.00%	16.54%	0.00%	0.00%	13.74%
Deep red pummelo	20.91%	20.76%	22.80%	21.68%	22.20%	18.81%	21.52%	26.04%	24.13%	21.91%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	14.02%	0.00%	0.00%	14.27%
KaoPan pummelo	12.64%	12.64%	10.44%	10.62%	11.21%	10.61%	9.79%	10.93%	12.26%	11.42%	0.00%	2.47%	0.00%	0.00%	0.00%	0.00%	17.56%	0.00%	0.00%	13.93%
Pink pummelo	8.75%	7.04%	5.78%	7.94%	9.40%	8.19%	9.55%	10.44%	7.63%	7.99%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	14.97%	0.00%	0.00%	13.69%
Tahiti pummelo	30.71%	31.75%	30.75%	31.69%	32.18%	28.31%	33.98%	29.85%	32.34%	31.33%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	14.65%	0.00%	0.00%	14.58%
Timor pummelo	4.54%	4.74%	3.95%	4.58%	9.04%	5.03%	7.26%	5.90%	6.27%	5.49%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	16.82%	0.00%	0.00%	13.82%
Chios mandarin	1.38%	1.44%	0.99%	1.47%	0.51%	0.84%	1.45%	0.98%	1.25%	1.08%	0.49%	0.51%	19.33%	8.94%	0.15%	44.88%	15.71%	33.00%	0.00%	81.30%
Cleopatra mandarin	1.78%	1.44%	2.07%	0.86%	3.40%	2.23%	3.26%	3.32%	1.74%	2.17%	0.25%	0.15%	5.71%	0.44%	1.27%	0.19%	11.75%	0.00%	0.20%	85.15%
Dancy mandarin	12.16%	12.72%	14.25%	14.08%	12.58%	11.92%	14.63%	12.53%	12.55%	12.89%	0.09%	30.12%	17.49%	0.20%	0.00%	20.61%	10.91%	30.06%	0.11%	77.95%
Fuzhu mandarin	2.51%	4.38%	4.28%	3.80%	4.27%	2.98%	5.32%	5.28%	4.83%	4.02%	0.00%	31.25%	18.81%	0.09%	0.00%	21.59%	12.64%	31.52%	0.00%	77.71%
King mandarin	20.83%	21.55%	20.97%	21.68%	18.94%	18.72%	23.70%	22.48%	22.97%	20.61%	2.35%	40.11%	47.17%	14.88%	2.23%	74.91%	29.00%	77.02%	46.99%	52.11%
Ladu mandarin	3.16%	5.82%	4.00%	4.75%	3.83%	5.59%	3.99%	3.93%	5.41%	4.37%	0.25%	8.92%	9.50%	0.27%	0.23%	0.00%	16.37%	14.45%	0.20%	82.95%
Ponkan mandarin	4.46%	5.03%	5.83%	5.44%	5.28%	5.68%	5.44%	6.27%	5.02%	5.25%	0.08%	23.83%	17.97%	0.18%	0.61%	44.42%	14.71%	18.74%	27.03%	74.89%
San Hu Hong Chu mandarin	2.51%	2.59%	3.57%	1.81%	2.68%	2.42%	4.11%	4.05%	2.32%	2.84%	0.25%	9.22%	0.05%	0.53%	0.74%	0.10%	15.26%	18.69%	29.74%	81.43%
Satsuma Owari mandarin	1.62%	3.09%	1.32%	1.99%	2.39%	3.07%	0.97%	2.33%	2.32%	2.07%	8.40%	14.16%	64.84%	0.18%	58.74%	15.47%	48.47%	54.34%	14.53%	51.73%
Sunki mandarin	12.56%	9.48%	12.36%	9.84%	12.80%	12.38%	11.25%	12.29%	9.27%	11.33%	0.37%	0.24%	7.24%	0.19%	0.17%	0.21%	15.53%	0.00%	0.64%	84.67%
Szibat mandarin	9.64%	9.34%	8.84%	9.59%	9.47%	10.61%	10.04%	13.64%	8.69%	9.74%	0.09%	29.87%	0.26%	3.72%	0.08%	0.00%	15.99%	9.39%	0.11%	82.09%
Carvalhal mandarin hybrid	2.67%	3.81%	2.96%	4.40%	4.41%	4.56%	3.39%	3.19%	3.28%	3.66%	0.67%	32.64%	49.81%	56.01%	48.87%	35.22%	13.77%	33.63%	30.84%	57.38%
Clemenules clementine	0.89%	1.65%	0.56%	1.12%	1.81%	0.65%	1.81%	3.44%	2.22%	1.38%	18.56%	32.94%	35.98%	30.13%	49.41%	45.64%	21.55%	29.26%	39.78%	63.82%
Dweet tangor	21.47%	21.55%	21.96%	22.11%	23.28%	21.42%	23.58%	24.82%	24.42%	22.36%	25.70%	59.25%	39.82%	82.93%	46.56%	24.53%	36.08%	51.96%	8.81%	44.93%
Ellendale tangor	12.64%	14.58%	14.25%	13.47%	17.14%	17.88%	16.69%	22.48%	17.18%	15.60%	30.33%	28.43%	19.30%	0.70%	48.34%	0.23%	29.90%	27.10%	45.22%	60.99%
Kiyomi tangor	16.21%	17.74%	16.74%	18.91%	14.46%	16.48%	24.06%	18.80%	15.64%	16.85%	30.46%	53.36%	63.35%	51.33%	64.84%	25.64%	37.58%	29.80%	47.03%	39.05%
Murcott tangor	18.56%	21.26%	20.03%	17.27%	18.80%	21.32%	21.52%	23.96%	18.24%	19.81%	2.39%	12.86%	17.11%	16.18%	3.29%	39.53%	13.87%	82.88%	37.90%	63.56%
Ortanique tangor	11.91%	8.26%	10.63%	10.45%	9.98%	9.96%	8.71%	9.95%	8.88%	9.73%	28.06%	42.21%	30.46%	59.50%	44.74%	78.80%	51.66%	30.42%	40.47%	41.74%
Temple tangor	22.12%	26.44%	22.38%	24.70%	22.42%	24.49%	27.69%	29.24%	21.04%	23.74%	19.67%	32.91%	48.33%	32.34%	24.88%	70.78%	19.40%	22.57%	39.00%	57.45%
Allspice tangelo	15.56%	14.80%	15.23%	13.56%	15.40%	14.43%	15.11%	16.83%	13.71%	14.76%	15.74%	75.55%	78.31%	83.72%	37.18%	85.20%	53.56%	29.10%	71.48%	26.43%

Annexe 7 : Analyse de l'hétérozygotie observée et estimation des données manquantes chez les variétés étudiées au niveau de chaque chromosome (Suite)

	Données manquantes										Ho									
	Chr1	chr2	Chr3	Chr4	Chr5	Chr6	Chr7	Chr8	Chr9	Total	Chr1	chr2	Chr3	Chr4	Chr5	Chr6	Chr7	Chr8	Chr9	Total
Mapo tangelo	35.01%	40.45%	38.41%	39.38%	34.56%	39.48%	39.30%	43.37%	35.62%	37.89%	45.14%	70.45%	48.78%	33.48%	33.48%	22.31%	41.83%	61.17%	38.38%	41.66%
Minneola tangelo	14.75%	14.30%	15.66%	15.63%	14.53%	15.27%	15.84%	15.23%	15.64%	15.04%	0.10%	73.43%	44.87%	40.63%	47.12%	37.25%	54.02%	78.99%	37.99%	41.98%
Orlando tangelo	8.02%	7.83%	9.40%	10.19%	9.91%	7.54%	7.98%	11.79%	8.20%	8.96%	10.31%	40.06%	37.57%	51.25%	48.48%	60.32%	51.38%	45.96%	26.92%	39.37%
Pearl tangelo	12.24%	11.93%	11.00%	14.16%	13.74%	14.71%	13.54%	11.92%	12.84%	12.53%	1.85%	77.08%	59.90%	84.31%	79.80%	85.04%	53.85%	44.07%	57.03%	27.51%
Sampson tangelo	11.83%	11.06%	11.52%	10.19%	10.77%	12.85%	13.91%	11.79%	10.42%	11.35%	0.83%	52.18%	58.18%	59.71%	45.38%	23.50%	38.06%	65.04%	62.07%	38.75%
San Jacinto tangelo	6.56%	4.74%	4.80%	4.58%	4.05%	5.68%	5.08%	6.14%	4.92%	5.12%	25.76%	77.30%	31.21%	24.34%	37.23%	37.02%	36.56%	49.74%	26.80%	48.95%
Seminole tangelo	18.96%	20.47%	19.42%	17.10%	22.27%	16.95%	19.71%	18.06%	18.73%	19.10%	5.80%	74.25%	71.30%	54.06%	54.79%	10.54%	52.41%	70.61%	38.95%	37.19%
Sunrise tangelo	20.18%	21.34%	23.93%	21.59%	23.36%	23.46%	23.58%	22.60%	21.04%	22.36%	54.01%	49.13%	27.26%	24.01%	29.34%	9.61%	9.02%	20.63%	0.37%	53.20%
Sunshine tangelo	41.65%	45.98%	44.48%	46.11%	45.77%	46.93%	44.86%	46.31%	44.59%	44.97%	77.50%	50.80%	39.88%	50.16%	41.33%	32.11%	35.09%	62.70%	44.43%	36.41%
UGLI® tangelo	30.39%	28.45%	30.18%	28.32%	29.07%	31.84%	33.37%	33.78%	28.38%	29.85%	64.61%	24.50%	62.83%	70.96%	39.14%	66.26%	29.76%	56.03%	31.81%	35.65%
Webber SG tangelo	12.07%	10.06%	7.95%	9.67%	10.05%	10.43%	10.16%	10.20%	8.11%	9.54%	69.68%	58.07%	65.07%	56.02%	40.43%	21.52%	21.94%	46.24%	41.28%	36.70%
Fairchild tangelo	2.27%	1.94%	2.30%	1.90%	2.53%	1.86%	3.02%	4.30%	2.90%	2.42%	3.73%	65.35%	32.29%	29.14%	39.09%	40.70%	24.19%	50.45%	30.02%	54.83%
Nova tangelo	11.91%	15.45%	13.02%	11.92%	12.44%	12.85%	17.65%	13.88%	12.55%	13.03%	25.11%	14.95%	38.54%	31.86%	56.73%	45.83%	53.16%	45.79%	29.58%	53.42%
Osceola tangelo	2.59%	3.02%	2.49%	1.12%	4.34%	2.14%	2.66%	4.30%	1.64%	2.67%	10.23%	27.70%	28.30%	49.08%	14.51%	8.18%	21.86%	30.55%	37.10%	56.39%
Page tangelo	1.30%	1.94%	2.26%	1.04%	3.62%	2.51%	5.20%	5.90%	2.32%	2.48%	0.74%	52.60%	51.32%	28.53%	54.54%	38.78%	28.44%	68.80%	38.93%	48.30%
Robinson tangelo	9.00%	12.79%	11.33%	10.45%	11.14%	12.66%	13.66%	14.99%	10.33%	11.38%	6.41%	36.49%	41.68%	22.95%	20.67%	0.21%	17.65%	35.40%	0.22%	62.07%
Fallglo mandarin	15.56%	16.81%	18.29%	19.34%	19.45%	18.72%	19.95%	21.13%	15.54%	17.87%	18.62%	16.15%	22.44%	36.62%	50.45%	40.78%	19.79%	73.99%	19.66%	57.59%
Kara mandarin	2.03%	1.22%	2.30%	1.30%	1.74%	1.58%	4.23%	1.72%	2.32%	1.74%	0.99%	40.58%	48.60%	0.35%	14.79%	30.18%	35.61%	78.38%	22.92%	57.61%
Wilking mandarin	12.80%	9.12%	11.66%	10.10%	9.83%	10.89%	10.88%	13.39%	11.00%	10.90%	1.58%	33.52%	24.43%	2.02%	0.96%	74.50%	14.11%	17.59%	28.31%	67.77%
Jackson orangelo	65.96%	68.32%	69.49%	69.52%	71.80%	70.11%	72.31%	67.94%	69.02%	69.21%	80.24%	18.59%	73.04%	84.70%	53.85%	1.56%	49.34%	54.79%	54.21%	14.09%
Triumph orangelo	15.72%	15.88%	15.89%	16.23%	18.58%	16.57%	19.47%	16.58%	17.08%	16.45%	86.25%	14.86%	79.88%	88.14%	58.79%	3.91%	51.50%	65.68%	56.69%	12.07%
Wekiwa SG tangelo	4.86%	7.11%	2.87%	5.87%	6.58%	7.26%	4.11%	4.05%	5.60%	5.11%	16.70%	29.31%	68.54%	33.21%	70.67%	83.94%	21.44%	65.94%	47.03%	32.94%
Fortune mandarin	1.05%	1.58%	1.46%	0.78%	2.60%	1.12%	1.81%	4.30%	1.64%	1.70%	6.72%	53.14%	40.60%	33.51%	48.55%	38.61%	42.36%	29.91%	10.40%	55.45%
Fremont mandarin	3.65%	3.52%	3.15%	4.92%	4.56%	3.63%	5.56%	5.53%	3.57%	3.89%	2.02%	32.84%	54.47%	1.27%	37.50%	45.51%	13.32%	38.23%	47.95%	61.55%
Ambersweet tangor	15.88%	15.23%	16.50%	14.77%	17.50%	15.18%	19.23%	18.18%	19.50%	16.51%	19.65%	44.32%	44.88%	72.24%	24.80%	49.07%	44.31%	59.31%	37.17%	36.06%
Marsh grapefruit	8.67%	11.49%	7.05%	6.30%	8.10%	7.54%	6.17%	9.71%	10.42%	8.27%	85.63%	13.39%	60.39%	65.71%	49.25%	84.49%	51.80%	66.39%	54.63%	7.43%
Star Ruby grapefruit	2.35%	4.09%	1.60%	2.68%	3.76%	1.68%	1.33%	3.56%	3.86%	2.69%	89.38%	13.11%	62.49%	66.64%	49.36%	86.17%	53.68%	68.28%	56.83%	6.64%
Bigaradier Maroc sour orange	0.73%	0.93%	1.74%	1.64%	1.30%	0.74%	0.85%	4.05%	2.32%	1.64%	90.29%	83.68%	86.41%	86.30%	80.73%	87.15%	53.78%	79.90%	79.25%	11.14%
Washington Navel Sweet orange	15.88%	15.30%	15.37%	14.85%	17.72%	14.99%	15.96%	19.29%	16.89%	15.90%	33.33%	46.48%	59.22%	88.03%	73.20%	85.76%	55.68%	70.62%	57.96%	26.73%

Annexe 8 : Analyse de l'homozygotie et de l'hétérozygotie des taxons de références chez les variétés étudiées à partir des points diagnostiques de chaque chromosome

	% DP HOM M										% DP HOM P									
	Chr1	chr2	Chr3	Chr4	Chr5	Chr6	Chr7	Chr8	Chr9	Total	Chr1	chr2	Chr3	Chr4	Chr5	Chr6	Chr7	Chr8	Chr9	Total
Chandler pummelo	0.00%	0.23%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.03%	100%	97.67%	100%	100%	100%	100%	100%	100%	100%	99.69%
Deep red pummelo	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
KaoPan pummelo	0.00%	0.16%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	100%	97.37%	100%	100%	100%	100%	100%	100%	100%	99.66%
Pink pummelo	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Tahiti pummelo	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Timor pummelo	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Chios mandarin	99.51%	99.49%	79.87%	92.80%	99.85%	51.55%	99.78%	63.90%	100%	87.69%	0.00%	0.00%	0.81%	0.16%	0.00%	3.57%	0.00%	3.10%	0.00%	0.77%
Cleopatra mandarin	99.75%	99.85%	93.81%	99.85%	98.73%	99.81%	99.89%	100%	100%	98.51%	0.00%	0.00%	0.48%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.10%
Dancy mandarin	99.91%	67.57%	80.48%	99.80%	100%	77.91%	100%	67.13%	100%	87.01%	0.00%	2.30%	2.03%	0.00%	0.00%	1.48%	0.00%	2.81%	0.00%	1.08%
Fuzhu mandarin	100%	66.34%	80.16%	100%	100%	76.39%	99.89%	67.57%	100%	86.75%	0.00%	2.40%	1.03%	0.00%	0.00%	2.02%	0.00%	0.91%	0.00%	0.80%
King mandarin	97.65%	55.13%	48.25%	87.65%	96.88%	19.70%	47.69%	15.06%	88.65%	60.44%	0.00%	4.76%	4.58%	0.49%	0.89%	5.38%	4.74%	7.92%	11.35%	4.04%
Ladu mandarin	99.75%	90.92%	89.03%	99.50%	99.77%	100%	99.55%	84.14%	100%	95.27%	0.00%	0.15%	1.47%	0.00%	0.00%	0.00%	0.00%	1.41%	0.00%	0.42%
Ponkan mandarin	99.92%	64.67%	80.48%	99.66%	99.39%	52.81%	99.32%	79.95%	96.10%	84.32%	0.00%	11.50%	1.55%	0.00%	0.00%	2.76%	0.00%	1.31%	3.90%	2.55%
San Hu Hong Chu mandarin	99.75%	90.04%	99.95%	99.84%	99.26%	99.90%	99.89%	78.10%	99.72%	96.75%	0.00%	0.74%	0.00%	0.00%	0.00%	0.00%	0.00%	3.20%	0.28%	0.37%
Satsuma Owari mandarin	91.52%	85.03%	30.25%	99.84%	38.52%	83.09%	13.55%	27.80%	99.42%	59.63%	0.08%	0.82%	4.91%	0.00%	2.74%	1.44%	3.96%	17.86%	0.58%	3.41%
Sunki mandarin	99.63%	99.76%	92.38%	99.82%	99.83%	99.79%	99.88%	100%	100%	98.34%	0.00%	0.00%	0.38%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.07%
Szibat mandarin	99.91%	67.51%	99.74%	97.06%	99.92%	99.90%	100%	89.33%	100%	94.67%	0.00%	2.61%	0.00%	0.00%	0.00%	0.10%	0.00%	1.28%	0.00%	0.45%
Carvalhal mandarin hybrid	99.25%	65.72%	38.71%	43.00%	49.55%	62.54%	100%	52.03%	99.42%	64.59%	0.08%	1.64%	11.48%	3.95%	1.59%	2.24%	0.00%	14.34%	0.58%	4.49%
Clemenules clementine	81.03%	65.30%	62.84%	72.47%	49.48%	51.08%	77.21%	68.07%	99.18%	66.85%	0.41%	1.75%	1.18%	1.56%	1.10%	3.28%	0.88%	2.67%	0.82%	1.47%
Dweet tangor	72.34%	36.90%	37.11%	23.92%	50.99%	74.53%	35.64%	31.54%	99.58%	51.30%	1.96%	3.85%	23.07%	7.42%	2.45%	0.95%	3.74%	16.50%	0.42%	7.91%
Ellendale tangor	67.63%	67.03%	67.05%	99.21%	45.55%	99.77%	60.61%	66.56%	91.70%	70.26%	2.04%	4.54%	13.65%	0.00%	6.11%	0.00%	1.52%	6.34%	8.30%	5.71%
Kiyomi tangor	68.28%	35.90%	32.02%	49.01%	22.99%	69.57%	34.75%	17.55%	74.95%	42.18%	1.26%	10.74%	4.63%	3.09%	12.17%	4.79%	17.90%	52.65%	25.05%	12.18%
Murcott tangor	97.21%	65.51%	66.84%	81.25%	96.62%	57.63%	99.59%	11.15%	93.16%	74.83%	0.40%	21.62%	16.05%	1.94%	0.09%	2.84%	0.00%	5.98%	6.84%	7.65%
Ortanique tangor	68.63%	50.67%	54.97%	34.45%	51.24%	12.72%	16.51%	63.44%	93.77%	49.27%	3.31%	7.13%	14.57%	15.64%	4.02%	8.48%	10.12%	6.14%	6.23%	8.58%
Temple tangor	79.50%	63.09%	47.79%	72.56%	72.23%	24.91%	81.13%	68.75%	95.59%	63.77%	0.83%	4.00%	3.88%	2.05%	2.89%	4.32%	0.28%	8.68%	4.41%	3.39%
Allspice tangelo	83.69%	18.97%	7.99%	18.13%	60.77%	9.79%	14.89%	58.64%	81.96%	34.22%	0.58%	5.48%	13.70%	8.17%	2.05%	5.01%	4.51%	12.26%	18.04%	7.11%

Annexe 8 : Analyse de l'homozygotie mandarine et pamplemousse des taxons de références chez les variétés étudiées à partir des points diagnostiques de chaque chromosome (Suite)

	% DP HOM M										% DP HOM P									
	Chr1	chr2	Chr3	Chr4	Chr5	Chr6	Chr7	Chr8	Chr9	Total	Chr1	chr2	Chr3	Chr4	Chr5	Chr6	Chr7	Chr8	Chr9	Total
Mapo tangelo	51.37%	22.80%	38.32%	65.68%	63.65%	75.08%	36.50%	11.71%	96.84%	48.36%	3.49%	6.76%	12.90%	4.43%	2.87%	2.62%	4.41%	27.11%	3.16%	7.58%
Minneola tangelo	99.62%	22.05%	52.06%	56.34%	51.52%	60.33%	10.66%	15.22%	95.02%	50.72%	0.29%	4.53%	3.07%	4.16%	1.35%	2.42%	3.68%	5.80%	4.98%	3.08%
Orlando tangelo	89.25%	19.41%	55.27%	46.01%	48.56%	34.54%	13.29%	49.86%	95.25%	49.63%	0.44%	40.53%	7.16%	5.19%	2.97%	5.14%	5.34%	4.18%	4.75%	9.46%
Pearl tangelo	97.88%	17.62%	26.10%	18.04%	14.67%	11.24%	12.04%	27.89%	95.36%	32.15%	0.28%	5.30%	14.00%	8.22%	5.53%	3.71%	5.53%	28.03%	4.64%	8.45%
Sampson tangelo	98.90%	7.84%	38.95%	41.28%	51.94%	74.79%	48.88%	12.26%	83.24%	48.27%	0.28%	39.98%	2.87%	4.77%	2.67%	1.71%	3.72%	22.70%	16.76%	9.99%
San Jacinto tangelo	73.46%	17.19%	51.36%	72.61%	61.12%	36.62%	35.22%	40.84%	96.39%	51.52%	0.78%	5.51%	17.43%	2.64%	1.66%	26.36%	4.50%	9.42%	3.61%	8.95%
Seminole tangelo	93.80%	22.04%	21.30%	44.56%	44.28%	89.13%	12.53%	23.54%	97.86%	46.18%	0.40%	3.70%	7.41%	1.71%	0.93%	0.34%	3.20%	5.85%	2.14%	3.26%
Sunrise tangelo	35.33%	10.68%	61.68%	75.54%	38.11%	89.90%	97.03%	62.22%	99.63%	59.04%	10.66%	40.18%	11.06%	2.40%	32.55%	0.49%	0.28%	17.14%	0.37%	14.67%
Sunshine tangelo	15.97%	11.30%	53.77%	48.28%	56.67%	65.09%	40.08%	6.64%	95.30%	41.68%	6.53%	37.90%	6.35%	11.21%	2.00%	2.81%	5.01%	30.66%	4.70%	11.66%
UGLI® tangelo	23.86%	67.27%	24.18%	25.91%	53.31%	24.86%	39.63%	14.29%	91.90%	39.84%	11.53%	8.23%	13.00%	20.06%	7.54%	8.88%	10.06%	29.68%	8.10%	11.98%
Webber SG tangelo	24.52%	8.55%	27.73%	40.73%	56.19%	75.88%	73.63%	13.54%	91.23%	41.36%	5.81%	33.39%	7.20%	10.46%	3.38%	2.60%	1.43%	40.22%	8.77%	12.00%
Fairchild tangelo	96.19%	31.21%	66.84%	69.61%	60.24%	57.21%	63.94%	45.19%	99.72%	64.17%	0.08%	3.44%	0.87%	6.30%	0.67%	2.09%	8.17%	4.36%	0.28%	2.44%
Nova tangelo	74.70%	60.32%	51.35%	73.32%	40.21%	51.50%	15.72%	45.51%	96.39%	54.76%	0.18%	24.72%	10.11%	0.96%	3.06%	2.67%	5.03%	8.70%	3.61%	7.53%
Osceola tangelo	89.60%	51.41%	70.59%	54.12%	58.65%	91.25%	77.29%	67.91%	94.38%	71.74%	0.17%	20.89%	1.11%	2.64%	26.83%	0.57%	0.34%	1.54%	5.62%	7.39%
Page tangelo	98.93%	22.86%	38.72%	73.45%	41.19%	58.26%	55.38%	27.28%	98.06%	52.83%	0.33%	24.54%	9.96%	2.95%	4.28%	2.96%	3.25%	3.92%	1.94%	7.27%
Robinson tangelo	92.97%	24.05%	54.35%	80.45%	76.73%	99.68%	79.98%	60.12%	100%	71.21%	0.62%	39.46%	3.98%	3.35%	2.60%	0.11%	2.20%	4.48%	0.00%	7.06%
Fallglo mandarin	79.85%	65.46%	66.74%	65.38%	47.40%	57.39%	76.82%	7.94%	98.01%	63.66%	1.54%	18.39%	10.82%	4.30%	2.15%	1.83%	0.26%	18.07%	1.99%	7.17%
Kara mandarin	98.68%	58.18%	47.88%	99.84%	84.69%	66.89%	33.96%	14.50%	98.21%	65.42%	0.33%	1.24%	3.51%	0.00%	0.52%	2.93%	3.64%	7.13%	1.79%	2.31%
Wilking mandarin	98.42%	64.19%	63.12%	97.78%	99.04%	22.05%	99.15%	80.00%	96.07%	77.00%	0.00%	2.29%	12.45%	0.18%	0.00%	3.45%	0.00%	2.41%	3.93%	3.71%
Jackson orangelo	12.86%	1.13%	9.71%	13.10%	11.54%	0.31%	13.24%	12.26%	14.29%	8.98%	6.90%	80.27%	17.26%	14.29%	34.62%	98.13%	5.51%	32.95%	85.71%	39.66%
Triumph orangelo	9.71%	1.45%	5.76%	16.60%	8.35%	0.45%	10.09%	5.15%	16.67%	6.88%	4.04%	83.69%	14.37%	2.98%	32.86%	95.65%	2.59%	29.16%	83.33%	36.65%
Wekiwa SG tangelo	82.28%	16.32%	27.88%	37.01%	11.84%	10.44%	71.73%	8.58%	93.63%	35.63%	1.02%	54.37%	3.58%	31.24%	17.49%	5.62%	8.38%	25.48%	6.37%	16.26%
Fortune mandarin	93.20%	44.31%	57.73%	65.53%	50.26%	50.47%	43.02%	68.04%	99.12%	62.06%	0.08%	2.55%	1.67%	6.65%	1.19%	10.92%	3.22%	2.05%	0.88%	2.89%
Fremont mandarin	97.98%	65.97%	41.80%	98.66%	60.68%	52.37%	99.55%	56.83%	93.65%	68.43%	0.00%	1.19%	3.74%	0.00%	1.82%	2.13%	0.00%	4.94%	6.35%	2.16%
Ambersweet tangor	60.69%	16.69%	35.92%	19.55%	63.63%	36.11%	9.83%	13.21%	98.28%	38.80%	19.65%	38.98%	19.20%	17.72%	11.57%	14.82%	25.16%	27.48%	1.72%	20.53%
Marsh grapefruit	10.38%	1.87%	4.75%	5.99%	6.45%	11.38%	11.10%	7.35%	16.39%	7.42%	3.99%	84.74%	34.85%	34.76%	44.30%	4.13%	7.98%	26.26%	83.61%	34.75%
Star Ruby grapefruit	7.72%	2.10%	3.54%	5.76%	6.69%	9.09%	9.49%	6.24%	12.33%	6.18%	2.90%	84.79%	33.97%	32.80%	43.95%	4.73%	4.75%	25.48%	87.67%	34.18%
Bigaradier Maroc sour orange	5.55%	10.37%	9.28%	17.58%	14.43%	9.10%	11.52%	11.65%	79.05%	12.14%	4.16%	5.95%	4.31%	2.98%	4.84%	3.75%	6.49%	8.45%	20.95%	5.35%
Washington Navel Sweet orange	65.61%	15.35%	37.44%	16.03%	21.53%	8.65%	11.62%	15.07%	91.44%	29.39%	1.06%	38.17%	3.33%	5.01%	5.27%	5.59%	3.03%	14.31%	8.56%	9.62%

Annexe 9 : Analyse de l'hétérozygotie des taxons de références chez les variétés étudiées à partir des points diagnostiques de chaque chromosome

	% DP He									
	Chr1	chr2	Chr3	Chr4	Chr5	Chr6	Chr7	Chr8	Chr9	Total
Chandler pummelo	0.00%	2.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.28%
Deep red pummelo	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
KaoPan pummelo	0.00%	2.47%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.32%
Pink pummelo	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Tahiti pummelo	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Timor pummelo	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Chios mandarin	0.49%	0.51%	19.33%	7.04%	0.15%	44.88%	0.22%	33.00%	0.00%	11.54%
Cleopatra mandarin	0.25%	0.15%	5.71%	0.15%	1.27%	0.19%	0.11%	0.00%	0.00%	1.40%
Dancy mandarin	0.09%	30.12%	17.49%	0.20%	0.00%	20.61%	0.00%	30.06%	0.00%	11.92%
Fuzhu mandarin	0.00%	31.25%	18.81%	0.00%	0.00%	21.59%	0.11%	31.52%	0.00%	12.45%
King mandarin	2.35%	40.11%	47.17%	11.85%	2.23%	74.91%	47.56%	77.02%	0.00%	35.52%
Ladu mandarin	0.25%	8.92%	9.50%	0.50%	0.23%	0.00%	0.45%	14.45%	0.00%	4.31%
Ponkan mandarin	0.08%	23.83%	17.97%	0.34%	0.61%	44.42%	0.68%	18.74%	0.00%	13.14%
San Hu Hong Chu mandarin	0.25%	9.22%	0.05%	0.16%	0.74%	0.10%	0.11%	18.69%	0.00%	2.88%
Satsuma Owari mandarin	8.40%	14.16%	64.84%	0.16%	58.74%	15.47%	82.49%	54.34%	0.00%	36.96%
Sunki mandarin	0.37%	0.24%	7.24%	0.18%	0.17%	0.21%	0.12%	0.00%	0.00%	1.58%
Szibat mandarin	0.09%	29.87%	0.26%	2.94%	0.08%	0.00%	0.00%	9.39%	0.00%	4.88%
Carvalhal mandarin hybrid	0.67%	32.64%	49.81%	53.05%	48.87%	35.22%	0.00%	33.63%	0.00%	30.93%
Clemenules clementine	18.56%	32.94%	35.98%	25.97%	49.41%	45.64%	21.90%	29.26%	0.00%	31.68%
Dweet tangor	25.70%	59.25%	39.82%	68.66%	46.56%	24.53%	60.61%	51.96%	0.00%	40.79%
Ellendale tangor	30.33%	28.43%	19.30%	0.79%	48.34%	0.23%	37.87%	27.10%	0.00%	24.03%
Kiyomi tangor	30.46%	53.36%	63.35%	47.90%	64.84%	25.64%	47.35%	29.80%	0.00%	45.64%
Murcott tangor	2.39%	12.86%	17.11%	16.81%	3.29%	39.53%	0.41%	82.88%	0.00%	17.52%
Ortanique tangor	28.06%	42.21%	30.46%	49.91%	44.74%	78.80%	73.37%	30.42%	0.00%	42.14%
Temple tangor	19.67%	32.91%	48.33%	25.38%	24.88%	70.78%	18.58%	22.57%	0.00%	32.83%
Allspice tangelo	15.74%	75.55%	78.31%	73.71%	37.18%	85.20%	80.60%	29.10%	0.00%	58.67%

Annexe 9 : Analyse de l'hétérozygotie des taxons de références chez les variétés étudiées à partir des points diagnostiques de chaque chromosome (*Suite*)

	% DP He									
	Chr1	chr2	Chr3	Chr4	Chr5	Chr6	Chr7	Chr8	Chr9	Total
Mapo tangelo	45.14%	70.45%	48.78%	29.89%	33.48%	22.31%	59.08%	61.17%	0.00%	44.06%
Minneola tangelo	0.10%	73.43%	44.87%	39.50%	47.12%	37.25%	85.66%	78.99%	0.00%	46.20%
Orlando tangelo	10.31%	40.06%	37.57%	48.79%	48.48%	60.32%	81.38%	45.96%	0.00%	40.91%
Pearl tangelo	1.85%	77.08%	59.90%	73.75%	79.80%	85.04%	82.43%	44.07%	0.00%	59.40%
Sampson tangelo	0.83%	52.18%	58.18%	53.94%	45.38%	23.50%	47.39%	65.04%	0.00%	41.74%
San Jacinto tangelo	25.76%	77.30%	31.21%	24.75%	37.23%	37.02%	60.28%	49.74%	0.00%	39.54%
Seminole tangelo	5.80%	74.25%	71.30%	53.73%	54.79%	10.54%	84.27%	70.61%	0.00%	50.56%
Sunrise tangelo	54.01%	49.13%	27.26%	22.06%	29.34%	9.61%	2.68%	20.63%	0.00%	26.28%
Sunshine tangelo	77.50%	50.80%	39.88%	40.52%	41.33%	32.11%	54.91%	62.70%	0.00%	46.66%
UGLI® tangelo	64.61%	24.50%	62.83%	54.04%	39.14%	66.26%	50.31%	56.03%	0.00%	48.18%
Webber SG tangelo	69.68%	58.07%	65.07%	48.81%	40.43%	21.52%	24.94%	46.24%	0.00%	46.64%
Fairchild tangelo	3.73%	65.35%	32.29%	24.09%	39.09%	40.70%	27.88%	50.45%	0.00%	33.38%
Nova tangelo	25.11%	14.95%	38.54%	25.72%	56.73%	45.83%	79.25%	45.79%	0.00%	37.72%
Osceola tangelo	10.23%	27.70%	28.30%	43.23%	14.51%	8.18%	22.37%	30.55%	0.00%	20.87%
Page tangelo	0.74%	52.60%	51.32%	23.60%	54.54%	38.78%	41.37%	68.80%	0.00%	39.90%
Robinson tangelo	6.41%	36.49%	41.68%	16.20%	20.67%	0.21%	17.83%	35.40%	0.00%	21.73%
Fallglo mandarin	18.62%	16.15%	22.44%	30.32%	50.45%	40.78%	22.92%	73.99%	0.00%	29.17%
Kara mandarin	0.99%	40.58%	48.60%	0.16%	14.79%	30.18%	62.40%	78.38%	0.00%	32.27%
Wilking mandarin	1.58%	33.52%	24.43%	2.03%	0.96%	74.50%	0.85%	17.59%	0.00%	19.29%
Jackson orangelo	80.24%	18.59%	73.04%	72.62%	53.85%	1.56%	81.25%	54.79%	0.00%	51.36%
Triumph orangelo	86.25%	14.86%	79.88%	80.43%	58.79%	3.91%	87.32%	65.68%	0.00%	56.47%
Wekiwa SG tangelo	16.70%	29.31%	68.54%	31.75%	70.67%	83.94%	19.89%	65.94%	0.00%	48.10%
Fortune mandarin	6.72%	53.14%	40.60%	27.82%	48.55%	38.61%	53.77%	29.91%	0.00%	35.05%
Fremont mandarin	2.02%	32.84%	54.47%	1.34%	37.50%	45.51%	0.45%	38.23%	0.00%	29.41%
Ambersweet tangor	19.65%	44.32%	44.88%	62.73%	24.80%	49.07%	65.01%	59.31%	0.00%	40.67%
Marsh grapefruit	85.63%	13.39%	60.39%	59.25%	49.25%	84.49%	80.92%	66.39%	0.00%	57.83%
Star Ruby grapefruit	89.38%	13.11%	62.49%	61.44%	49.36%	86.17%	85.76%	68.28%	0.00%	59.64%
Bigaradier Maroc sour orange	90.29%	83.68%	86.41%	79.43%	80.73%	87.15%	81.99%	79.90%	0.00%	82.51%
Washington Navel Sweet orange	33.33%	46.48%	59.22%	78.96%	73.20%	85.76%	85.35%	70.62%	0.00%	60.99%

Résumé

L'étude de l'espaceur intergénique *trnL-trnF* chez les espèces tunisiennes du genre *Citrus* montre la présence d'une seule copie du pseudogène *trnF* chez toutes les variétés analysées. Les valeurs positives et non significatives des tests de neutralité plaident en faveur d'un modèle d'évolution neutre et supportent l'hypothèse d'un scénario démographique stable. Nos résultats montrent clairement la contribution des pools génétiques de *C. reticulata* et *C. maxima* dans le développement des espèces secondaires tunisiennes. L'analyse basée sur les séquences nucléotidiques de huit régions génomiques chloroplastiques de 79 variétés de la sous famille des Aurantioideae révèle, via les marqueurs SNPs, une différenciation taxonomique au niveau des tribus, des sous-tribus, des genres et des espèces. 166 SNPs diagnostiques des 54 clades analysés ont été identifiés suivis d'une sélection de 40 KASPar. L'application de ces marqueurs chez 108 variétés montre un haut taux de transférabilité au sein de la sous famille des Aurantioideae et une cohérence avec les analyses génétiques antérieures du génome chloroplastique. 21 clés marqueurs de clade catégoriquement diagnostique de 19 clades ont été identifiées. L'analyse GBS des 55 variétés d'agrumes par séquençage *ILLUMINA HISEQ2000* s'avère efficace dans l'identification des polymorphismes diagnostiques (12564 SNPs/InDels) de la différenciation *C. reticulata/C. maxima* couvrant l'ensemble du génome nécessaire au déchiffrement de l'origine phylogénomique tout au long des neuf chromosomes des 55 variétés. L'approche adoptée confirme les analyses récentes basées sur des données de séquence de génome complet pour la clémentine, l'orange douce et amère et la mandarine 'Ponkan'. La technique GBS couplée à la détection des points polymorphes diagnostiques s'avère très efficace dans le décryptage des caryotypes phylogénomiques des variétés qui dérivent d'une mosaïque de deux espèces ancestrales. Le mélange *C. reticulata/C. maxima* doit être l'élément majeur de la variabilité phénotypique révélée élevée de ces ressources.

Mots clés : *Citrus*, Aurantioideae, espaceur intergénique *trnL-trnF*, régions chloroplastiques, SNP, InDel, KASPar, GBS, séquençage, diversité génétique, origine génétique, phylogénie, structure interspécifique.

Abstract

The study of the *trnL-trnF* intergenic spacer in Tunisian *Citrus* species show the presence of a single copy of the pseudogen *trnF* in all the analysed varieties. The positive and non-significant values of the neutrality tests plead In favour of a neutral evolution model and support the hypothesis of a stable demographic scenario. Our results shows clearly the contribution of genetic pools of *C. reticulata* and *C. maxima* in the development of Tunisian secondary species. Analysis based on the nucleotide sequences of eight chloroplast genomic regions of 79 varieties of the Aurantioideae subfamily reveals SNPs taxonomic differentiation at tribes, sub-tribes, genera and species levels. 166 diagnostic SNPs of the 54 analysed clades were identified, followed by a selection of 40 KASPar. The application of these markers on 108 varieties shows a high transferability rate within the Aurantioideae subfamily showing coherence with previous chloroplast genetic analyses. 21 clade markers categorically diagnosed from 19 clades were identified. The GBS analysis of the 55 *Citrus* varieties via the *ILLUMINA HISEQ2000* sequencing proved to be effective in identifying the diagnostic polymorphisms (12564 SNPs/InDels) of the *C. reticulata /C. maxima* differentiation covering the whole genome necessary to decipher the phylogenomic origin throughout the nine chromosomes. The approach adopted confirms recent analyses based on complete genome sequence data for clementine, oranges and mandarin 'Ponkan'. The GBS technique coupled with the detection of diagnostic polymorphic points is effective in decrypting phylogenomic karyotypes of varieties derived from a mosaic of two ancestral species. The mixture *C. reticulata / C. Maxima* should be the major factor in the high phenotypic variability of these resources.

Key words: *Citrus*, Aurantioideae, *trnL-trnF* intergenic spacer, chloroplast regions, SNP, InDel, KASPar, GBS, sequencing, genetic diversity, genetic origin, phylogeny, interspecific structure